Universidade de Aveiro Departamento de Electrónica, Telecomunicações e Informática,
2018

João Paulo Aires de Matos Melo

Towards a profile of emotions: a Psychophysiological data mining proof of concept

De encontro a um perfil psicofisiológico das emoções: uma prova de conceito baseado em data mining

Universidade de Aveiro Departamento de Electrónica, Telecomunicações e Informática,
2018

João Paulo Aires de Matos Melo

Towards a profile of emotions: a Psychophysiological data mining proof of concept

De encontro a um perfil psicofisiológico das emoções: uma prova de conceito baseado em data mining

Dissertação apresentada à Universidade de Aveiro para cumprimento dos requisitos necessários à obtenção do grau de Mestre em Engenharia de Computadores e Telemática, realizada sob a orientação científica da Doutora Susana Manuela Martinho dos Santos Baía Brás, Investigadora do Departamento de Eletrónica Telecomunicações e Informática da Universidade de Aveiro e do Professor Doutor José Maria Amaral Fernandes, Professor auxiliar do Departamento de Eletrónica Telecomunicações e Informática da Universidade de Aveiro.

o júri / the jury

presidente / president

Professor Doutor Tomás António Mendes Oliveira e Silva
Professor Associado do Departamento de Eletrónica, Telecomunicações e Informática da Universidade de Aveiro (por delegação da Reitora da Universidade de Aveiro)

vogais / examiners committee

Professor Doutor Miguel Tavares Coimbra
Professor Auxiliar do Departamento de Ciências de Computadores da Faculdade de Ciências da Universidade do Porto

Doutora Susana Manuela Martinho dos Santos Baía Brás
Investigadora do Instituto de Engenharia Eletrónica e Telemática de Aveiro / Departamento de Eletrónica , Telecomunicações e Informática da Universidade de Aveiro

agradecimentos / acknowledgements

Aproveito a oportunidade para agradecer a todas as pessoas que me seguiram ao longos destes árduos meses de trabalho, que direta ou indiretamente estiveram sempre presentes com o apoio que conseguiam disponibilizar.

Em primeiro aos meus caros amigos que sempre me apoiaram nos maiores momentos de stress, aqui sublinho o nome de Tiago Henriques, que esteve sempre presente, nos almoços, nas árduas idas ao ginásio para relaxar, na constante companhia e risos, nas noitadas no IEETA sempre com boa música, nas boleias no Híbrido dele, mas mais importante agradeço o constante apoio, também um especial obrigado ao Alexandre Oliveira e Filipe Tavares, por me terem aturado ao longo desta jornada, com apoio tanto sentimental como técnico e nas jogatanas de Lol, amigos destes nunca serão esquecidos. Aos meus amigos de Viseu, Pedro Batista por me fornecer sempre conversas do melhor que há, estar sempre a disposto a aturar-me e de me fazer evoluir o meu gosto musical, também à Inês Peres por estar presente há muitos anos e ser uma das melhores pessoas que já conheci, obrigado aos dois por me aturarem mesmo quando estou a chorar, aos outros meus melhores amigos Alexandre Corte-Real, Henrique Santos, Carolina Gonçalves e Sérgio Lopes que me aturaram sempre fora desta atmosfera e me fizeram relaxar e acalmar, um grande obrigado a todos. Um grande agradecimento ao resto do grupo de Aveiro que mesmo que não tenham estado presentes tão vivamente nesta fase foram das partes mais importantes deste percurso académico como um todo.

Em segundo aos meus orientadores por me fornecerem sempre um ambiente de confiança, e por terem tido o trabalho de ler este extenso trabalho. À Doutora Susana Brás por fornecer conhecimento científico e técnico nas variadas áreas que por mim não eram dominadas, à constante paciência mesmo quando eu demorava a entender determinado assunto, por me ter dado uma enorme vontade de escolher este tópico devido à paixão que me passou a quando da apresentação deste e por além de realizar trabalho de orientadora e professora também realizar trabalho de amiga em momentos de menor motivação, deixo um especial agradecimento com o maior carinho. Ao professor José Maria Fernandes por ter sido sempre um grande professor ao longo destes anos, por sempre me ter ensinado como fazer diagramas que mesmo hoje ainda não sei fazer e por criar um bom ambiente de convivência.

Aos colegas de Psicologia, Margarida Alves, Catarina Costa, Inês Silva, Beatriz Sousa, Beatriz Ribeiro, Stelya Pereira, Lígia Ribeiro, Leonor Pereira e Sofia Fernandes, que realizaram a experiência da qual os dados para comparação foram adquiridos, a todos os participantes nas variadas experiências, e em especial à Filipa Barros que esteve sempre perto a dar feedback e a acompanhar o trabalho. Também à Professora Sandra Soares e Marta Rocha que me forneceram dados e explicações cruciais para a criação do sistema de identificação.

À minha mãe, por toda a dedicação que introduziu na minha educação, não só a nível académico, mas também por me fornecer sempre com os melhores princípios que irei levar para sempre no coração, o maior dos obrigados por teres sido a mãe mais perfeita de sempre, penso que nunca te irei conseguir agradecer tudo o que fizeste por mim, sendo que a única coisa que posso fazer é ser a pessoa mais exemplar que conseguirei. Por todas as vezes que me deste carinho e acreditas-te em mim, por todas as chamadas diárias só para saberes como estava, por toda a confiança que me deste, mas especialmente por todo o amor que me forneceste. Foste a pessoa que mais me apoio desde sempre e sei que irás ser a que mais irá apoiar até ao fim, sempre serás quem me irá levantar e me dar forças, mesmo que não as tenhas arranjas só para me dar. Obrigado do fundo do coração, nunca te esqueças que és o meu ídolo, a minha pessoa favorita e a pessoa que mais irei amar neste mundo. Obrigado. Também à minha avó por ser a minha segunda mãe e ter estado presente em todos os anos da minha vida, obrigado por tudo.A toda a minha família por me aturar e estar sempre presente, sei que são uma cambada de malucos, mas são a minha família e tenho um enorme orgulho de pertencer a esta família. Obrigado a todos.

Por último, mas não esquecidos o resto das pessoas que sempre me apoiaram e acreditaram em mim, mesmo quando o único que queria era desistir, desde amigos a familiares, aqui não terei que nomear nomes porque cada um deles sabe quem é.

A todos, obrigado por me terem ajudado nesta fase da vida.

Keywords Aprendizagem Automática; Extracção de dados; Sistema Portátil de monitorização; Psicofisiologia; ECG; HR; EDA; EMG

Resumo A capacidade de reconhecer, compreender e controlar emoções é um dos aspetos mais importantes da inteligência humana. Dar esta capacidade a máquinas pode ser um grande benefício para a humanidade, sendo que com isto o sistema pode facilmente adaptar-se às necessidades do utilizador baseado no seu estado emocional, com esta adaptação melhorias na vida diária do utilizador podem ser alcançada. Neste trabalho nós desenvolvemos prova de conceito que é capaz de integrar informação de sensores vestíveis e identificar emoções, que podem, mais tarde, ser usadas em sistemas nomeadamente para adaptar a aplicação e/ou o ambiente. Utilizando este tipo de sistemas em pessoas com maiores necessidades como idosos, crianças, pessoas com distúrbios mentais, entre outras, poderão vir a fornecer um enorme melhoramento nas tarefas diárias mais básicas, assim como aprimorar a sua independência. Nós aplicámos *machine learning* para a classificação de três emoções humana – Alegria, Neutra e Medo. O sistema consegue identificar o estado emocional baseando-se em quatro sinais fisiológicos: Eletrocardiografia (ECG), Eletromiografia (EMG), Atividade eletrodérmica (EDA), e Respiração (RSP). O bloco de integração, este irá adquirir os sinais fisiológicos fornecidos pelos sensores, o bloco de identificação, que irá ser responsável por classificar a emoção baseando-se no sinal fisiológico como input e um bloco final responsável por fornecer feedback baseando-se na classificação da emoção. Antes da classificação, selecionamos features a partir dos sinais disponíveis usando várias técnicas nomeadamente, *feature variance threshold*, *correlation threshold* e *principal componente analysis* foi também usada para redução de *features*. Para a fase de classificação diferentes métodos populares foram usados, *K-nearest neighbours* que atuava com uma precisão de 60% e *neural network* que apresentava uma precisão de 50% na classificação de emoções. Os sinais fisiológicos que nós usamos não são ideais para monitorizar pessoas e gravar as suas respostas corporais a estímulos emocionais em rotinas diárias. Por essa razão, como parte do trabalho nós também realizámos uma avaliação do Vital Jacket (sensor vestível) contra o sistema *Biopac* (*gold standard* em configurações de laboratórios) em ambos o ECG e o Batimento Cardíaco. Para o ECG, foi observado que existe um atraso aleatório que faz a comparação direta difícil (Correlação *Spearman* de 0.44 com uma janela de 5 segundos). Para o batimento cardíaco o erro foi quase 0. Apesar do reconhecimento emocional estar ainda nos estágios iniciais, este trabalho mostra que um sistema de reconhecimento emocional é possível e viável e poderá ser considerada a sua utilização na vida real usando uma solução vestível com as considerações apropriadas nas limitações conhecidas e de acordo com o âmbito da aplicação.

Keywords Machine Learning; Data Mining; Portable Monitoring systems; Psychophysiology; ECG; HR; EDA; EMG

Abstract The capacity to recognise, employ, comprehend and manage emotions is one of the most important aspects of human intelligence. Giving this capacity to machines can be a huge benefit for mankind since with it, the system can easily adapt to the necessities of the user based on their emotional state. In follow up, responsive machines can be created, and with them an improvement in one's daily life can be achieved. In this work we developed an end-to-end proof of concept solution that is able to integrate information from wearable sensors and identify emotion, that can, later on, be used by other systems namely to adapt application and/or environment. We applied machine learning for the classification of three different human emotions - Happy, Neutral and Fear. The system can recognise the human affective state based on four physiological signals – Electrocardiography (ECG), Electromyography (EMG), Electrodermal activity (EDA) and Respiration (RSP). The system is divided into three blocks. An integration block that will acquire data from sensors, an identification block responsible to classify the emotion based on physiological inputs and a final block responsible to provide feedback based on the classification. Prior to the classification, we selected features from the available inputs using several techniques namely variance threshold, correlation threshold. Principal Component Analysis was also used for feature reduction. For the classification phase two different popular methods were used, k-nearest neighbours that performed with an accuracy of 60% and neural networks that yield an accuracy of 50% on emotion classification when the features did not pass through PCA and no features from EMG were used. The physiological signals we used although a way to monitor persons and record their body response to emotional stimuli might not be ideal in real daily-life routines. For that reason, as part of the work we also performed an assessment on Vital Jacket (wearable sensor) against Biopac system (gold standard in laboratory settings) in both ECG and Heart Rate (HR) For the ECG, random delays were observed making the direct comparison difficult (Spearman correlation of 0.44 in 5 second windows). For HR the error was near 0. Although emotional recognition is still in early stages, this work shows that an emotional recognition system is possible and viable and be considered to use in real life using wearable solution with proper considerations on known limitations and on the scope of the application.

Contents

List of Figures

List of Tables

List of Acronyms

ASD Autism spectrum disorder

ECG Electrocardiography

EMG Electromyography

EDA Electrodermal activity

RSP Respiration

VJ Vital Jacket

Std Standard Deviation

Knn K-Nearest Neighbors

PCA Principal component analysis

Chapter 1

Introduction

People engagement on daily life routines is increasingly relevant for a healthy society [1]. Anxiety disorders affect more than 450 million sufferers worldwide [2]. Being anxiety one of the most prevalent mental disorder, it is crucial to find solutions that may help people dealing with anxiety. Which is even more relevant to individuals with incapacities, their families and the community around them, since severe symptoms of anxiety substantially worsen their life. Therefore, innovative solutions for the investigation and management of anxiety is a priority.

Emotion regulation is important in our well-being, and decision making. It is important to feel fear, because our mind and body learns how to deal with extreme or dangerous situations, however, the misunderstanding of this emotion overloads our body. Also, the emotional regulation or deregulation is associate with immune responses [3], which affects the person health condition, and his/her ability to respond to virus/ bacteria and diseases. Emotion identification, therefore, plays an important helping people in their regulations. Anxiety as an example appears from an imbalance or difficulty to regulate emotions [4].

Human centred perspective is gaining a momentum [5], and computer science community is now giving a way to use its tools in order to think in the human and society as the central factor [6]. Now, we know that the human perspective plays a major role in every system. So, identifying the person feelings will allow to improve the interaction. When we identify emotion, we are able to describe the person, and therefore his preferences. The affective computing is defined by the study and development of systems able to recognise, interpret, process and simulate human affects. The major idea is to modulate the environment, e.g., an adaptive game, a robot that interact with a kid, a smart home. Identifying an emotion will allow to understand the person and adapts the environment to his/her needs.

Physiological response of emotions is studied as a mean to share emotions with others [7]. Several factors are involved in the emotional response, face recognition, posture analysis, physiological signal processing (electrocardiogram, electroencephalogram, respiration, electrodermal activity, electromyography). Industry is now developing technology that is able to overcome the needs to accomplish such goal. Sensors received a major evolution, in part caused by the evolution of technology, causing the reduction of the price, improvement of the specifications and reducing the intrusiveness of them, and in part by the ”quantified self” movement going mainstream [8], all of this caused the sensors to pass from a mere ”tech gadget” to a ”daily life” need by everyone [9]. Although the principal goal is to interpret the person, all system should be less intrusive as possible. With the evolution of sensors, the decision support system also had a major evolution, in the past 40 there was a bigger necessity to synchronize the information upon update. These evolution was done in 3 dimensions, the first one was from reactive to prescriptive, the reporting of the early 60’s were meant to provide an idea of what happened, as opposed to today solutions that provide recommendations of future actions, the second one from strategic long-range to operational short-term, starting by using reports to understand long term trends to construct solutions to provide real-time feedback, and the third one from internal function-oriented to external integrated, that put aside the organization view of information to a more driven view that integrated across functions and organizations. This new vision can be summarized in Fig. 1.2.

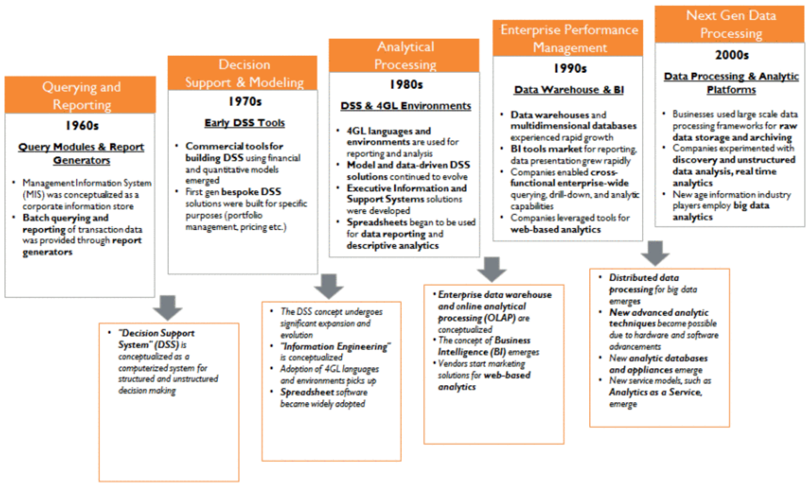


Figure 1.1: Summary of the key points in decision making systems
(Source: https://goo.gl/kPdHko)

Simultaneously with an associated growth in data there was a natural ever-growing focus in machine learning, that is defined as an advanced application of AI, that interconnects machines and peripherals granting them access to databases to make them learn new things from it and with that make them responsive [10], so, with so much more data, these algorithms can improve a lot more and more techniques are being found 1.2

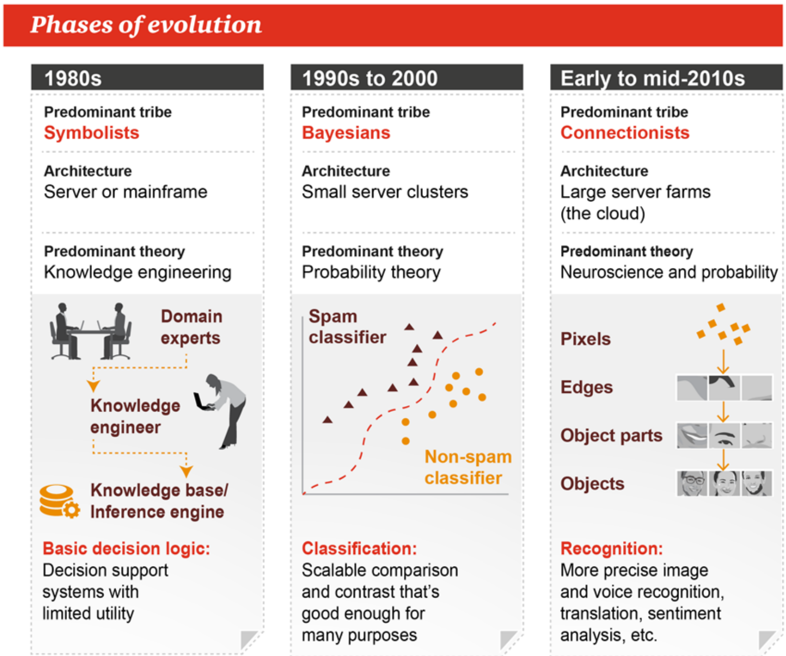


Figure 1.2: Phases of the evolution of machine learning
(Source: https://goo.gl/Bawfm5)

The study of human physiology is a good example of an application area where all these evolutions may merge to produce useful results namely in understanding emotions and suitable decision support to classify them. However, only the interdisciplinary relation between the areas can allow to understand the emotion and built a decision support system.

With all of this in mind an end-to-end proof of concept that will identify the human affect state and then it will return it as an emotion using a non-intrusive sensor as the collecting point is going to be presented in this work, with it a non-intrusive system will be created that can give almost real-time feedback to the user giving the possibility of improving one's daily life.

1.1 Motivation

Emotional communication is crucial for social interactions. Emotional communication is performed namely by visual, olfactory or muscular stimuli. However, there are persons with difficulty in the emotional interpretation. By the opposite, other persons do not communicate appropriately their feelings. A special case is persons with ASD. Automatic emotion identification systems are specially interest in those cases, since they can help in the emotional interpretation facilitating their social interactions.

Innovation in technology creates the opportunity to interpret and analyse individual data to prevent and advise the person. In this context emotion identification plays a key role, so, the development of an automatic emotion identification system will allow to validate and interpret the psycho-physiological profile of emotions, person individual characteristics and environmental emotional characteristics.

The major goal of this dissertation is the development of an end-to-end proof of concept solution that integrates information from wearable sensors, process it and identify an emotion, that can, later on, be used by other systems namely to adapt application and/or environment. The major idea is to apply the system in order to adapt to environments and communicate the emotions felt by the user.

1.2 Objectives

Regarding the goals and challenges presented, the major idea is to improve the state of the art in the emotional automatic identification system based on physiological data. So, to accomplish such goals, we proposed to:

- Deploy a physiological data mining and machine learning pipeline system.
- Evaluate machine learning models for emotion classification;
- Understand the physiological relation with emotions by means of feature selection;
- Evaluation of non-intrusive sensors;

1.3 Dissertation Outline

Excluding this one, 5 chapters compose this document:

- Chapter 2 - describes the state of the art associated to our system;
- Chapter 3 - Describes the system, showing the architecture, and a description of each block;
- Chapter 4 - Describes the identification block in detail, describing the setup and data selection protocol. Also, the data mining and machine learning techniques used and the emotion classification;
- Chapter 5 - Summarises the conclusions drawn from the development and testing stages, as well as a subsequent analysis;
- Appendices - In this appendix the comparison between the Vital Jacket and the Biopac MP160 is presented. A description of the data collection protocol is presented, and the results explained;

Chapter 2

State of the art

The scope to produce an emotion identification system is within the areas of Affective Computing and Psychology. In affective computing the interpretation, simulation and processing of emotions is increasingly evolving. The interdisciplinary concept inherent to this area facilitate the collaboration between psychology, computer science, informatics and electronics.

In this chapter it will be reviewed existing solutions for emotion identification system.

2.1 Emotion and emotional context

An emotion is described by William James [11] as ”the bodily changes follow directly the perception of the exciting fact, and that our feeling of the same changes as they occur is the emotion” [11].

Emotions are also characterised as subjective and not deterministic. Therefore, the same stimulus can reproduce totally different emotions in different individuals. However, the same stimulus in different days can recreate different reactions for the same individual [12].

Several models were created throughout the years to try to explain emotions, some of these models define emotions through a small number of primary and distinct emotions. Others suggest that emotions are better defined of as broad dimensions of experience (e.g., a dimension ranging from pleasant to unpleasant).[12] The first theory can be described as categorical approach and some of the names of the defenders of this theory are presented in the table 2.1 as well as the emotions they define.

Table 2.1: Theorists and their emotions

Theorist	**Basic Emotions**
Ekman, Friesen, & Ellsworth, 1972	Anger, disgust, fear, joy, sadness, surprise
Frijda, 1986	Desire, happiness, interest, surprise, wonder, sorrow
Gray, 1982	Rage, terror, anxiety, joy
Izard, 1977	Anger, contempt, disgust, distress, fear, guilt, interest, joy, shame, surprise
James, 1884	Fear, grief, love, rage
Mower, 1960	Pain, pleasure
Oatley & Johnson-Laird, 1987	Anger, disgust, anxiety, happiness, sadness
Plutchik, 1980	Anger, acceptance, joy, anticipation, fear, disgust, sadness, surprise
Tomkins, 1984	Anger, interest, contempt, disgust, fear, joy, shame, surprise

Emotions can be divided into four distinct components: Feelings; Bodily arousal; Sense of purpose; Social-expressive. Each of these components as multiple points:

- Feelings
  - Subjective experience
  - Phenomenological awareness
  - Cognitive interpretation
- Bodily arousal
  - Bodily preparation for action
  - Physiological activation
  - Motor responses
- Sense of purpose
  - Impulse to action
  - Goal-directed motivational state
  - Functional aspect to coping
- Social-expressive
  - Social communication
  - Facial expression
  - Vocal expression

With these components the understanding of emotion can improve in some way.

The first experiment that tries to understand emotions dated to 1872, with a work signed by Charles Darwin [13]. Darwin stated that ”The study of Expression is difficult, owing to the movements being often extremely slight, and of a fleeting nature.” ([13]), he believed that the display of facial expressions when an emotion occurred was an aid to survival, because this gave the capacity to people to communicate their internal states and react to problems even before the development of languages.

As indicated in the tables 2.2 and 2.3, there is growing evidence that emotional states can be mapped respectively to their corresponding specific physiological signals. [14] This table describes how several physiological signals of an individual vary when some specific emotion is being felt. It can be verified that some physiological variables vary based on the emotion being felt by the individual.

Table 2.2: Emotion to Signal relation

Signal	**Emotion**	**Result**	**Ref**
Facial EMG	Happiness / sadness / anger	EMG accurately discriminated between all conditions	[15]
Heart Rate / Skin conductance	Anger / fear / sadness / disgust / happiness	Anger, fear, and sadness increase the heart rate more than disgust. Anger and fear increase the heart rate more than happiness. Fear and disgust increase skin conductance more than happiness.	[16]
Skin conductance	Happiness / fear	The levels of tonic arousal and phasic skin conductance where higher on fear	[17]
Heart Rate	Neutral / fear	Heart rate was higher during fear imagery than neutral imagery	[18]
Facial EMG / Heart rate / Skin conductance	Pleasant emotional experiences	Eyebrow frown and smile are associated with the pleasantness dimension, Heart rate measure offered strong support between effort and arousal. Skin conductance offers further support but not as strong as heart rate	[19]

Table 2.3: Continued

Signal	**Emotion**	**Result**	**Ref**
Heart rate / Skin conductance / Facial EMG	Disgust / anger / pleasure / joy	Disgust, joy, and anger imageries offered a higher heart rate than pleasant imagery. Anger and disgust could be differentiated using facial EMG	[20]
Heart rate / skin conductance / finger temperature / blood pressure / EOG / Facial EMG	Neutrality / fear / joy / action / sadness / anger	An accuracy of 99% was achieved	[21]
GSR / heart rate / skin temperature	Attention / concentration / happiness / sadness / anger / fear / disgust / surprise / neutrality	No recognition found, some observations only	[22]

2.1.1 Emotion characterization

Physiology describes the function of organs, and how they react within the body to respond to external stimuli. Following a list of some physiological signals examples is presented:

- Heart Rate (HR) - is the number of beats the heart of a person performs per minute [23];
- Electrocardiography (ECG) - it is the electrical signal from the heart. It appears from the potential difference. It is recorded from the body surface; [24]

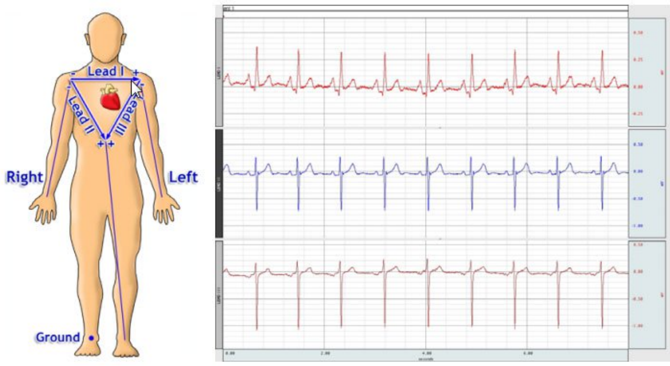


Figure 2.1: ECG sensor and signal
(source `https://bit.ly/2K9pjes`)

- Electromyography (EMG) - it is a signal that measures the electrical muscle response [25];

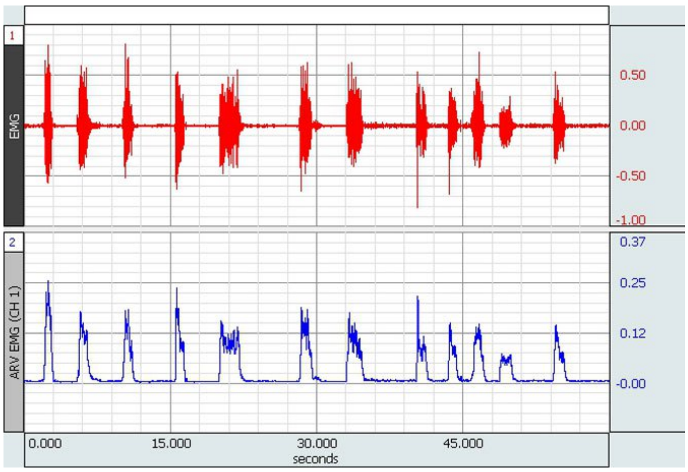


Figure 2.2: EMG Signal
(source `https://bit.ly/2KRPc3G`)

Figure 2.3: Facial Electromyography Signal
(source: `https://bit.ly/2I4uZty`)

- Electrodermal activity or Galvanic skin response (EDA/GSR) - it's a signal that represents the level of conductivity of electricity of the skin, the increase in this conductivity can occur by external or internal stimuli that are physiologically arousing [26];

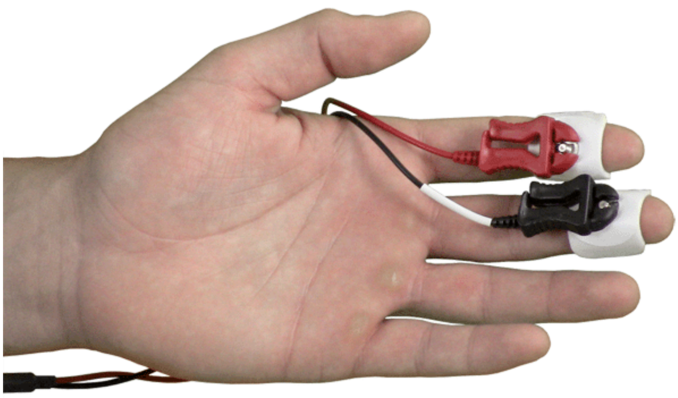


Figure 2.4: EDA sensor
(source `https://bit.ly/2rzr9yg`)

Emotions trigger alterations on central and peripheral nervous systems, those alterations communicate with human organs to adapt to the stimuli the person is facing [27]. Therefore, the response registered in the alteration in physiological signals may indicate the emotion and may help in the definition of a psycho-physiological response of emotion. [28]. There are also several evidences that human emotions are organized by neural system of appetitive and defensive motivation that mediate a range of attentional and action reflexes.

2.2 Affective computing

Affective Computing is a cross-disciplinary field from computer science, psychology, and cognitive science. [29]

The modern branch from computer science originated with Rosalin Picard with her book published in 1995 with the same name has the branch ”Affective Computing” [30], and it is described by her hands as ”computing that relates to, arises from, or deliberately influences emotions.” [31].

The research on this field manly focus on the design of systems that can be assigned with the human-like capabilities of interpretation and generation of affect features. [29] With this the quality of human-computer communication will be improved and thus improving also the intelligence of the computer [32].

The affective computing tries to build an ”affect model” based on several information gathered from multiple sensors (Video, Audio, Text, Vital Signs, etc), building then a personalized system that will give us intelligent, sensitive and friendly responses by having capability of perception, interpretation of the feelings of the human using it [32].

The ultimate purpose of affective computing is to ”teach” the computer how to properly react after the interpretation of the user's affect and meaning of that affect, not only react in this precise moment but also learn the changes of the user's affect. [32]. To give this personalized response the system needs to in some way gather external information and process it, so it can interpret what it's happening. The key technologies to process this information are:

- Emotional Speech Processing or Speech - this kind of information is transmitted through the explicit linguistic message (what is said) and the implicit features of the expression (how it is said), the last one has not yet been fully understood how the listeners can decode basic emotions through it [33], of course, this decoding is not only by voice but also other non-linguistic vocalisations. Affect detection based purely on speech has less accuracy than based on facial expressions, being that sadness, fear and anger are the emotions best recognised through the voice. But voice can be considered a promising signal due to his low-cost and non-intrusive resolution; [32] A work in this area was done by Chul Min Lee, using as source the information from language, discourse and acoustic and combinations of it. The selection of features was done by several methods like, all the features, the 10 best ones, the 15 best ones, with PCA, for the classification phase linear discriminant classifiers (LDC) with Gaussian class-conditional probability and k-nearest neighbours (k-NN) were used and an error rate between 10 and 30 percent was achieved depending on the method and on the information given. [34];

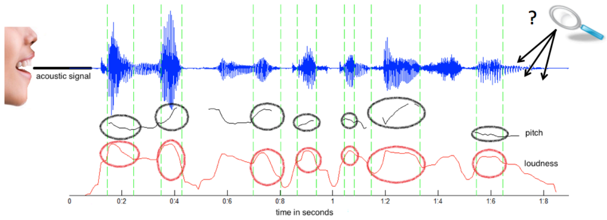


Figure 2.5: Extracting emotional relevant features from speech signal
(Source: `http://human-ist.unifr.ch/seminars/emotion-recognition`)

- Facial expression recognition and generation - There are several facial expressions and movements that can be used to fulfil some semantic functions, to communicate emotions and/or as conversational cues [32]. To better understand this Etcoff [35] created 37 lines to determinate the structure of the main parts of the human face, Ekman [36] then created a facial action coding system, with this, the facial expressions were deconstructed into action units and six basic emotions were described (joy, anger, surprise, disgust, fear and sadness). A work in this area was performed by Rana El Kaliouby [37] and, it was done the classification of complex states like agreeing, concentrating, disagreeing, interested, thinking and unsure. So, this was possible 25 fiducial landmarks were tracked for the head and facial detection using FaceTracker [1], facial actions are identified from motion, shape and colour descriptors derived from the feature points. The classification phase was created using a deep belief network (DBN) with an overall accuracy of 77.4%;

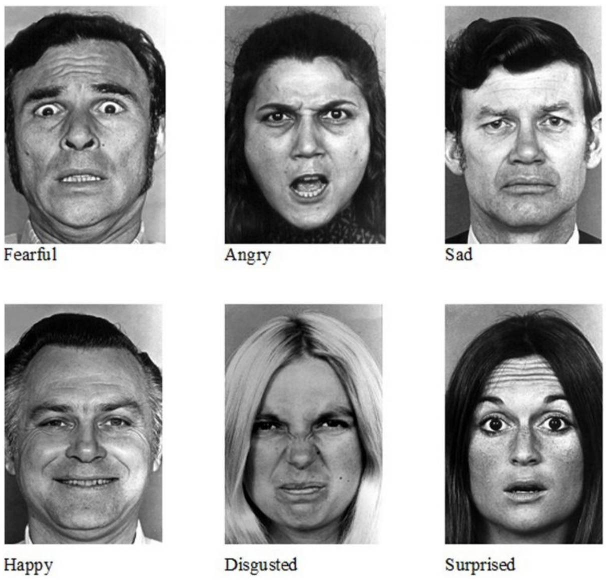


Figure 2.6: Examples of the Ekman-Friesen Pictures of Facial Affect
(Source: `https://www.frontiersin.org/articles/10.3389/fpsyg.2015.00761/full`)

[1] `https://github.com/lqs/neven`

- Body gesture and movement - this is defined as the position of body arthroses and their movement through time. These positions can be combined and temporally aligned, and with this a posture is created and a potentially ideal affective communicative channel is created [32]. Even though the primarily researches conducted by Darwin focused mainly on this topic state-of-the-art systems overlook this greatly compared to facial expressions and acoustic-proposed features [38]. Even if this gives information sometimes not available through the other two channels, an example of this is that an affection of a person can be deduced over long distances with posture, but the same is difficult with facial features. [39]. One of the works done in this area was by the hands of Hatice Gunes [40], in here he used body gestures to detect 6 emotions (Anxiety, Anger, Disgust, Fear, Happiness and Uncertainty) the feature extraction was done with the aid of that detected the movement and position of the body and the use of a BayesNet for the classification phase resulted in an overall accuracy of 94.02%;

Figure 2.7: Examples of body postures and their associated emotion (Source: `https://www.vision.ee.ethz.ch/publications/papers/articles/eth_biwi_00545.pdf`)

- Multimodal system - Normally the human to human interaction is done, by default with a multimodal interaction, this means, the participants encounter a set of facial expressions, gestures, body postures and movements, words, grammatical constructions, and prosodic contours to communicate and understand the other. [38] So, with this the main idea is to make an aggregation of multiple sensors. This can be done in three different forms: Data fusion - done on the raw signals that have the same temporal resolution, it is commonly used in the case of physiological signals[38]; Feature Fusion - this technique is done on the features extracted from the different signals, these are individually computed for each signal and combined across the sensors. [38]; Decision Fusion - this technique is done by classify each of the signal from the sensors and then merging the output of them all. [38] A work in this area was done by Soujanya Poria [41], here it was proposed a tri-model (text, audio and video) to detect 6 emotions (Surprise, Joy, Sadness, Anger, Fear and Disgust). Features for building training models were extracted from the data sets for each modality, then each of this feature vectors were classified by an SVM and the outputs of all were fused using the proposed feature-based achieving with this an accuracy of 87.95%;

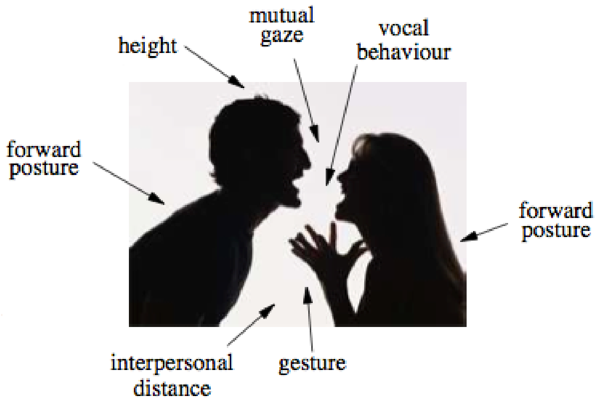


Figure 2.8: Multimodal system inputs examples
(Source: `http://human-ist.unifr.ch/seminars/emotion-recognition`)

- Text processing - this refers to the detection of emotions based on written language and transcriptions of oral communication. A work in this area was done by Wei-Hao Lin[42], it was used OpenNLP tools to automatically extract sentence boundaries, and reduced word variants using the Porter stemming algorithm to create the feature vector. SVM and naive Bayes were the classifiers used for the classification phase and got an accuracy of 85% and 87%;

- Physiology - focused on detecting affect with the help of machine learning techniques that will identify patterns in the physiological activity, this pattern will be correlated with the expression of different emotions. [32] In here several electrical signals produced by the body are recorded, being measured by Electromyogram, Electrodermal Activity, Electrocardiogram, Electrooculogram, Electroencephalography and others [32]. A work in this area was done by Changchun Liu [43], the main objective was to identify liking, anxiety and engage in children in the Autism Spectrum. The signals used here ECG, EMG, impedance cardiogram (ICG), photoplethysmogram (PPG), phonocardiogram (PCG), electrodermal activities (EDA), relevant features were derived from the physiological signals using various signal-processing techniques such as Fourier transform, wavelet transform, thresholding, and peak detection and with the use of an SVM a mean accuracy of 82.9% was acquired;

The existing researches are at large limited in detailed and scattered fields like voice and body language, and this is mainly caused by the lack of large affect data resources, no effective technique for multi-feature affective computing and the insufficiency of adaptation to natural scenarios [32]. With this computer cannot accurately judge and generate human-like affect status or have a real affect interaction. But, in recent years, the evolution of wearable computing, creates an opportunity to make computing technology a huger part on our daily life's, bringing with it a plus to the real time capture of affect information as well as providing a better platform for affective computing [32].

Figure 2.9: Affective Computing
(source : `https://vlab.org/events/affective-computing/`)

2.2.1 Automatic Emotion identification

Emotion identification involves the assessment of person signals and alterations. Emotion may be inferred by face expression alteration, posture, behaviour or physiological signals [44], [45], [45].

Kim [45] demonstrated how it was possible to classify three emotional statuses: sad, stressed and angry. Using ECG and EDA, features were extracted like, RR interval, heart rate variability, SCR, and using these features as input for a pattern classification method (support vector machine), the following percentage of correctly classified data points could be achieved: 89.7% for the training set and 55.2% for the test set.

Genki Okada in his paper [46], used an RGB camera to try and detect six emotions (amusement, anger, disgust, sadness, surprise and fear). Using pulse wave signal changes, several features were extracted from Heart Rate Variability and Facial Skin Blood Volume, being then used as input in a k-nearest neighbour classification method. The results gotten were the following: 94% accuracy with k=4 or k=5, while fear being excluded as this was the emotion that got worst results.

There are also other works like the work of Egon L. van den Broek [47], in his work he used several features extracted from physiological signals (EDA, ECG, EMG, skin temperature and respiration). For the feature selection and reduction phase it was used Analysis of Variance (it examines the variance of data set means compared to within class variance of the data sets themselves) and PCA. For the classification phase three different models were used SVM with average recognition rate of 60.71%, k-NN with k=8 and a rate of 61.31% and ANN with a rate of 56.19%.

There were also several works that tried to recognise emotions (sadness, happiness, anger and neutral state) outside of Physiological signals like the work of Carlos Busso [48]. In his work, in all different approaches it was used support vector machine as the classifier. For the speech system statistics of the pitch and the intensity were used and an overall performance on the classifier of 70.9% was achieved. For the facial expression approach the spatial data of the markers present on the face is reduced to a 4-dimensional feature vector using a PCA method, with this an accuracy of 85% was possible. A bimodal system was also proposed, this system fused the other two approaches into one and the performance of it was 89.1%, here the sadness got the worst results has it got confused with the neutral state.

As affective computing is an area that needs in some way to recognise the user emotion to fully use his capacities it needs tools to make this recognition. Different tools / sdk's / software's were created throughout the years [49] In order to aggregate signals and infer emotion, there is already available applications in the market. In Table 2.4 a summary of several applications is presented, by verifying this table it can be seen that different api's that detect emotions from different inputs already exist and can be used freely.

With the table 2.4 it can be seen that there are a lot of commercially available API's for emotion detection. They mainly use face recognition or linguistic analysis to make this emotional analysis. There are little to none API's that use physiological signals to understand emotions.

Table 2.4: Emotion Apps

Application	***What is it?***	***Main Features***	***Techniques***
Affectiva [a]	It is an API that analyses spontaneous facial expressions that people show in a daily basis, based on pictures, videos or even webcam live stream	Works without internet; Can be totally mobile / cloud; Works with live stream	Convolutional neural networks for multi-category, multi-label classification; Region proposal networks for fast object (face) detection and tracking; Recurrent neural networks for audio and video processing
Microsoft Emotion API [b]	The Emotion API takes a facial expression as input and returns the confidence across a set of emotions.	Synergy with Azure for cloud usage; It can perform near-real-time analysis on frames taken from a live video stream;	Information not provided
IBM Watson Tone Analyzer [c]	API to perform linguistic analysis to detect joy, fear, sadness, anger, analytical, confident and tentative tones present in the text.	Can detect at document and sentence level;	SVM to perform the predictions; One-Vs-Rest paradigm;
Kairos [d]	Commercial-grade emotion analysis, face detection and recognition engine provided as a public API. Kairos takes the complexity out of facial recognition and emotion analysis, so you can focus on building a great product.	Face Detection; Face Identification; Face Verification; Emotion Detection (joy, surprise, sadness, fear, anger, and disgust); Demographic, Feature Detection (age, gender, attention, dwell, glances, blinks, feature points, glasses, ethnicity); Multi-face tracking; Face Grouping.	Information not provided

[a] `https://affect.media.mit.edu`
[b] `https://azure.microsoft.com/en-us/services/cognitive-services/emotion/`
[c] `https://tone-analyzer-demo.ng.bluemix.net/`
[d] `https://www.kairos.com`

2.3 Autism spectrum disorder

This work is early proof of concept of a bigger project that has a main objective of dealing with children with autism spectrum disorder. Even though this work never passed to the phase of dealing with children with ASD, mainly because of these being individuals with needs, have difficult with dealing with strangers and strange systems. Nevertheless, the study of emotions and anxiety in children with ASD was done so the system was created based on this. For this reason, this chapter of the state of the art exists.

Autism spectrum disorder (ASD) is described as a developmental disorder. And, it is known that communication and behaviour is compromised, since persons with ASD does not interpret and understand the emotions in their peers, and, they cannot communicate their emotion [50]. Although this disorder may be diagnosed at any age, usually it is diagnosed in the first two years of the children, since it is a developmental disorder, signals appear earlier and allow paediatricians to earlier start intervention, which guarantees more success.

According to the Diagnostic and Statistical Manual of Mental Disorders (DSM-5) [51], some traces for people with ASD are presented:

- Difficulty with communication and interaction with others;
- Restricted interests and repetitive behaviours;
- Symptoms that hurt the person's ability to function properly in school, work, and other areas of life;

Figure 2.10: Autism Spectrum Disorder
(Source: `https://goo.gl/QwRCBk`)

As Autism is defined as a "spectrum" disorder this means that there is a wide variability in the type and severity of the different symptoms. Therefore, intervention and therapy should be customised to the person needs [52].

According to the Centre for Disease Control and Prevention (CDC) around 1 in 58 children has been identified with some form of ASD [53].

The Autism Spectrum can be divided into three severity levels, accordingly to the current diagnostic manual (DSM-5) [51].

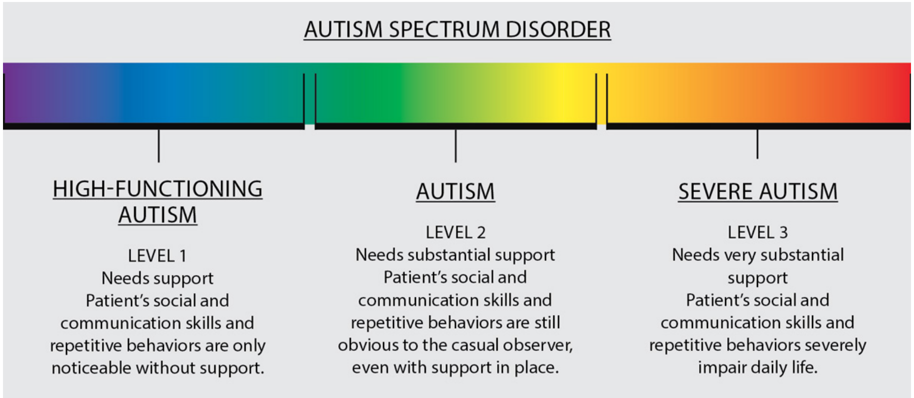


Figure 2.11: Autism Spectrum
(Source: `https://goo.gl/QwRCBk`)

The first level subtitled "Requiring Support" is the higher level in the spectrum and can be characterized as having difficulties in initiating social interactions, can present difficulty switching between activities, problems with organization and planning.

The second level subtitled "Requiring substantial support" is the middle level in the spectrum and is characterized as: having deficits in verbal and nonverbal communication skills; reduced response to social overtures from others; social impairments; restricted / repetitive behaviours appear frequently enough so they are susceptible to the casual viewer; Difficulty in changing focus; Inflexibility of behaviour.

The third and last level is subtitled "Requiring very substantial support" is the lower level in the spectrum and is characterized as: having several deficits in verbal and nonverbal communication skills; very limited initiation of social interactions; minimal response to social overtures from others; extreme difficulty coping with change; restricted / repetitive behaviours appear frequently enough to interfere with the functioning in all spheres; Inflexibility of behaviour; Great difficulty in changing focus.

2.3.1 Emotions and anxiety in Autism

Emotions

People with ASD may find difficulties in the understanding of the following social clues:

- Recognise facial expressions as well as the emotions associated to the expression [51];
- Mimic or use the facial expressions [51];
- Control the way they show emotions and understand them [51];

- Understand and interpret emotions, this can lead to lack, or seem to lack, of empathy from others [51];

Problems in the expression of emotions, usually appears around the age of school [54], at this age children with less severe ASD will exhibit almost the same emotions as the normally developed, however their explanation is hard [54]. On the other hand, children with more severe ASD will start displaying less emotions and empathy than the normally developed in an earlier stage [54].

Detecting emotions in individuals that suffer from autism can be a real challenge, several projects and papers have been done among the last years trying to recognise to try and understand this, one of this works was done by Changchun Liu [43], he used two computer-based cognitive tasks to elicit the affective states of liking, anxiety, and engagement, this state are considered important in autism intervention. The signals used in this paper were electrocardiogram (ECG), Impedance cardiogram (ICG), Photoplethysmograph (PPG), Phonocardiogram (PCG), Electrodermal activity (EDA) and Electromyography (EMG), acquired with the Biopac MP150, to extract features from this several techniques were used like Fourier transform, wavelet transform, thresholding, and peak detection. Due to the fact that autism is a spectrum disorder no intervention technique will work for the entire population, six participants in the age range of 13 to 16 years volunteered for this experiment (Each had a diagnosis on the autism spectrum, either autistic disorder, High-Functioning, or pervasive developmental disorder not otherwise specified). For the classification of the emotion an SVM was used and a correct prediction accuracy across all participants with ASD were 85.0% for liking, 79.5% for anxiety, and 84.3% for engagement.

Another work by the same author [55], he presented a novel affect-sensitive human-robot interaction framework for rehabilitation of children with ASD, with the aim of enabling the robot to, based on the affective cues of the children respond in order to help them explore social interaction. So, this was possible the following physiological signals were used: electrocardiogram (ECG); Photoplethysmograph (PPG); Impedance cardiogram (ICG); Bioelectrical impedance analysis (BIA); Electromyography (EMG). Three participants in the age range of 13 to 15 years volunteered for this experiment (Each had a diagnosis on the autism spectrum, either autistic disorder, High-Functioning, or pervasive developmental disorder not otherwise specified). For the classification phase an SVM with kernel RBF (Radial Basis Function) was used with an accuracy of approximately 83%.

Anxiety

Every children have to experience all emotions, positive and negative ones. Usually the more difficult to deal are negative ones, being sometimes a challenge from children and parents to understand rage, frustration and anxiety. ASD children suffer from episodes of high levels of anxiety [54], which has a negative impact on child behaviour [54]. The exteriorization of such negative feelings is according to ASD manifestation characteristics - stemming, obsessive, ritualistic behaviour and resistance to changes in routine [54].

Anticipation of anxiety episode may be critical to improve the daily life quality of children. Therefore, it is critical to find means to detect anxiety levels.

Several works to try and detect the anxiety in children with this kind of disorder have been done throughout the years, Hui-Chuan Chu in his paper develops a facial expression-based emotion recognition. Facial landmark tracking is performed using the FACEAPI to track multiple landmark points and four basic emotions (Calm, Happy, Anxious and Angry) are classified. Using Landmark transformation, Statistical processing (to measure the signal variance) and normalization a feature vector is got. The classifier used was an SVM and it was behind a recognition rate of 98.54% [56].

A project at Stanford Medicine named Autism Glass Project is being developed, they are developing a system that uses a combination of machine learning and artificial intelligence to automate facial expression recognition, this system is running on wearable glasses and can simultaneously provide social cues within the child's natural environment, it can also record the amount and type of eye contact, which can add more to the behavioural intervention. The expression recognition system is built on a light face tracker and the machine learning classifier achieves classification accuracy of 97% [57] [58].

Chapter 3

System

The main goal of the system is to create an end-to-end proof of concept, that would collect physiological data from wearable sensors, send it to an identification system that would respond with an output that will represent the class of the emotion present in the signal.

So, to accomplish such goal, the system that resulted from this work is divided into three blocks, an Integration Block that will record the data from a given sensor and send it to the second block, the second block is the Identification Block that will receive the information sent from the first block as input and through data mining and machine learning techniques it will classify the emotion, this classification will be sent to the Communication Block that will provide feedback based on the classification.

A full architecture of the end-to-end system can be checked on the figure 3.1.

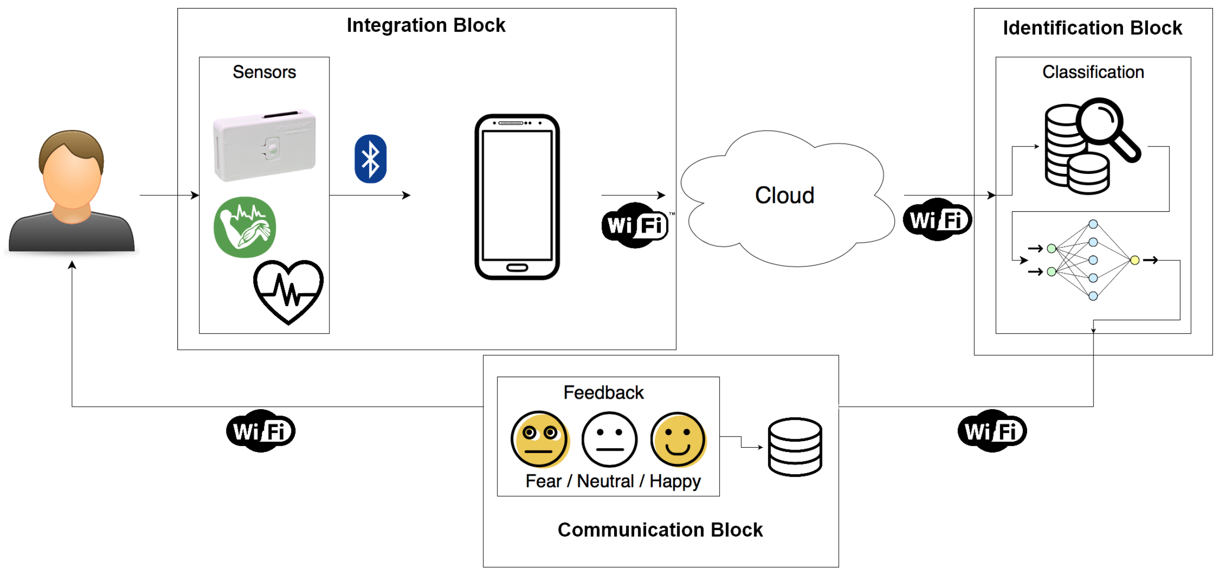


Figure 3.1: System Architecture

3.1 Integration Block

For the first block of the system the idea was to create a block that would make the communication between the gathering system (the sensor) and the identification system (the identification block). Recording the Vital Signals during the experiment and then creating the bridge to the next block.

Implementation

For the implementation of this block an android application was created that had as main goal of receiving data from any sensor, and store it, in this proof of concept it was used the Vital Responder as the sensor to record the physiological signal. With the reception of the data done from the sensors it will then be automatically sent to the cloud, so it can be used by the next block.

The details of this application are the following:

The application was designed in order to the user start by specifying the id of the participant and the session, and the session could now be started.

When the "Start" button is pressed the session will begin, the application try to connect to the Vital Jacket. If the connection is successful the session is started normally, if not the user can try to reconnect to it.

During the session there are multiple triggers that can be fired, these triggers are the "Start Baseline" and "Start Movies", they will activate baseline mode and experiment mode correspondingly. This will write the stimulus id to the files (0 - "Nothing", 1 - "Baseline", 2 - "Movies").

At the end, the session can be stopped and 4 files (ECG file, Accelerometer file, Heart Rate file and QRS Complex file) are created and can be sent directly to the cloud so they can be used by the next block with the press of a button.

As its main goal is to be used easily by any kind of user independently of their technological knowledge the user just needs to add the id of the participant and the session can easily start.

The main advantage of this app over the already created Vital Responder application called VRUnit is that for the purpose of this work the existing application provided too much features not necessary, so for sake of simplicity it was decided to implement a more specific application while reusing from VRUnit the needed features besides implementing the new ones.

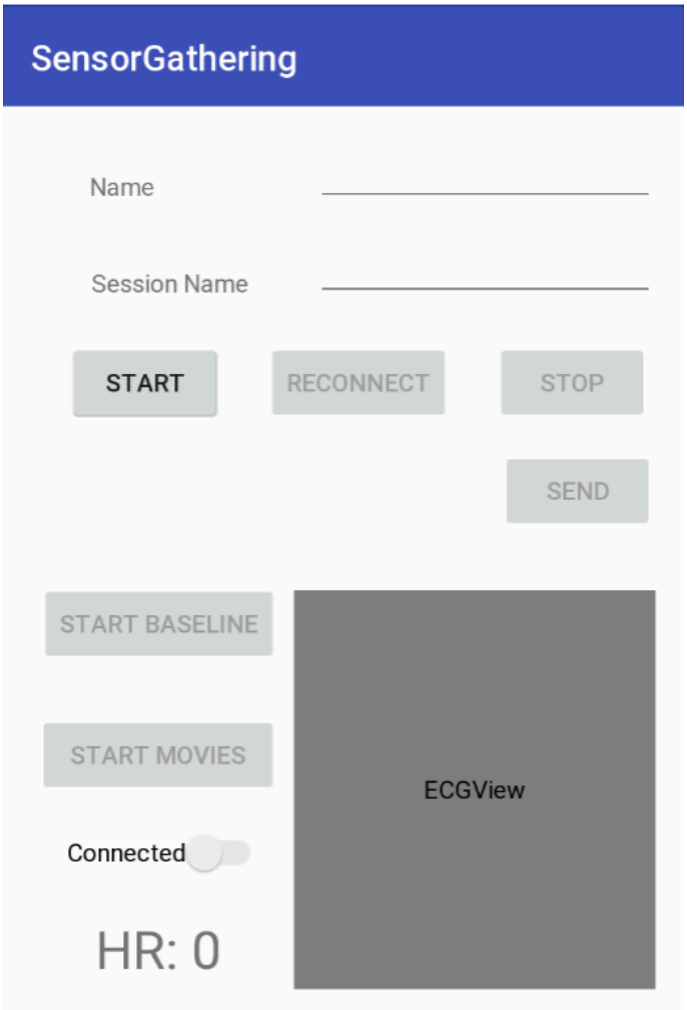


Figure 3.2: Application

3.2 Emotion identification block

The second block had the main objective of after receiving the information from the first one identifies the emotion present on this set of data.

Implementation

For the implementation of this block two parts were created, the first one is data mining, this part is responsible for extracting metrics from the sensor signal and, from these, extract the relevant features. The selection of both features and metrics is described in section 4.4.

The second part of the block is the classification module, the algorithm will provide emotion identification based on features extracted based on ML models. The selection of features and models is described in section 4.5.

Figure 3.3: Identification Block

As this block is the primary objective of this dissertation it will be explained in more detail later (chapter 4).

3.3 Communication Block

The third and last block had the purpose of after the identification was done from the previous block this would create a bridge for this result so that it could be used later.

Implementation

For the implementation phase of this it was created a network communication block, that will receive the prediction of the previous block, sent it to a cloud storing mechanism so it can be later used. So, the prediction is sent as a string that displays easily the emotion of the participant (”Fear”, ”Happiness”, ”Neutral”), to a Firebase [1] cloud database, where it will be store so it can be easily retrieved.

[1] https://firebase.google.com/

Chapter 4

Identification Block

In this chapter the implementation details of the main block are going to be explained, as well as, the details of the selection of the relevant information (metrics and features) in section 4.4 and the method to better classify the emotions based on them in section 4.5. So, this was possible, a specific experiment from emotions study was selected as departure point, described in section 4.1.

4.1 Case Study

The study of emotions, their physiological and behavioural answer is a theme that is frequently studied in the Department of Education and Psychology at the University of Aveiro. In 2015, a study was designed to access the emotional content of movies, and the physiological response of the induced emotions.

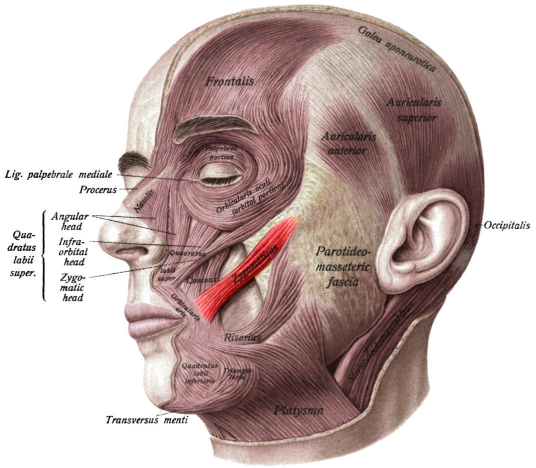
The study's participants were recruited between the university students. The participants were asked to fill a pre-selection questionnaire which addressed multiple variables, as: age, gender, health issue, actual medication. The inclusion criteria for the experiment comprised of the following: be a girl, heterosexual, Caucasian, non-smoker, healthy donors (medication free, without respiratory, metabolic or mental illness), normal or corrected to normal vision, score in the normal range of Trait Anxiety Inventory (STAI), aged between 18 and 35, not pregnant or suspected pregnancy, college students, native language (Portuguese), right handed, handedness scale, they should not be capable of understanding Mandarin.

There were 44 participants (all females) in this experiment with an age mean of 22.07 and a standard deviation of 3.7, and their age varied between the 18 and 35 years old.

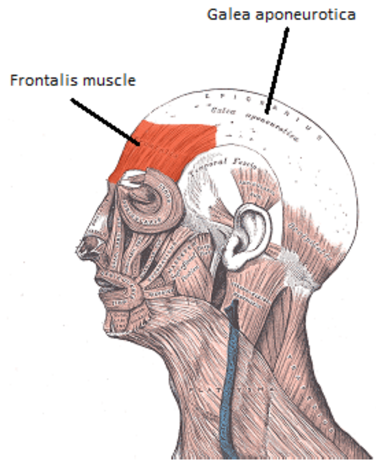
Data was collected in three different sessions per participant, where a different emotion was elicited. The first two sessions have no temporal separation, however the third one presented a separation of two hours from the end of the first two sessions.

4.2 Protocol and Setup

Participants wore multiple sensors from the Biopac system, this is a flexible, proven modular data acquisition and analysis system for life science research [59]. The sensors used collected ECG, EDA, Facial EMG (recorded at the Zygomatic Major muscle (4.1a) - associated with the movement for the smile, and at the Frontalis muscle (4.1b) - associated with the movement for fear [60]), and a band to record the respiration (RSP). All of the data was recorded with a sampling rate of 200 Hz.



(a) Zygomatic Major muscle location (Source: `https://goo.gl/XogMvD`)

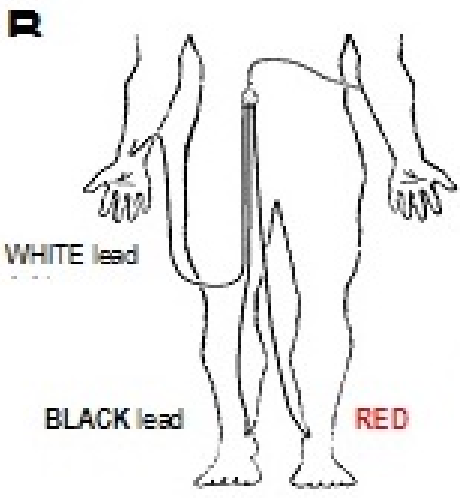


(b) Frontalis muscle location (Source: `https://goo.gl/tEFHzJ`)

Figure 4.1: EMG muscles

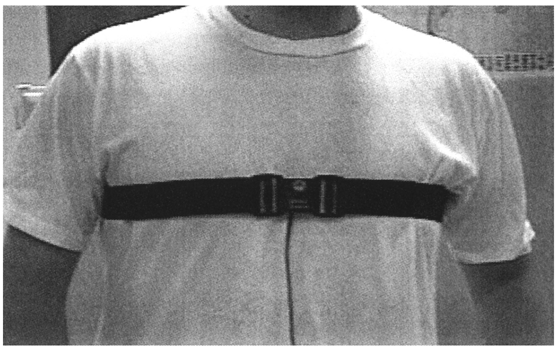
The electrodes positioning was as follow:

- ECG - lead two positioning fig. 4.2a. Negative electrode was placed at the right arm near the palm. Positive electrode was placed at the left leg near the foot, on the ankle. Ground electrode was placed at the right leg near the foot, on the ankle;
- EDA - positioned at the Medial Phalanx of the index and the middle finger of the left hand (fig. 4.2b);
- Respiration - the Biopac RSP band was positioned above the sternum (fig. 4.2c);
- EMG - the electrodes were positioned two on the Zygomatic Major muscle with a distance of 1 centimetre between each other; and two on the Frontalis muscle, one at the centre and the other with a distance of 1 centimetre and slightly towards the right side (fig. 4.2d);

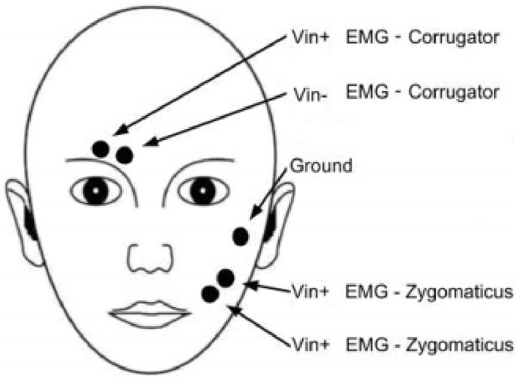


(a) Biopac sensors positions
(Source: `https://goo.gl/5guGox`)

(b) EDA electrodes positioning
(Source: `https://goo.gl/9efraZ`)



(c) RSP band positioning
(Source: `https://goo.gl/s9tqss`)



(d) Facial EMG electrodes positioning
(Source: `https://goo.gl/baSnSg`)

Figure 4.2: Electrodes positioning

The experience was divided in three phases were the participants faced three different emotions: fear, happiness, and neutral, by watching excerpts of movies. The order of the exposition of the participants to emotions was counterbalanced, that is, the participants would always watch the neutral clips first but the presentation order of the emotional clips (fear and the happy) will be counterbalanced across participants (e.g., participant 1: Neutral/Fear/Happy; participant 2: Neutral/Happy/Fear; participant 3: Neutral/Fear/Happy, etc.).

The protocol followed several steps:

1. After participants' selection, a first interview was scheduled. This interview was used to explain what the participant would do during the experiment, the different sensors he would wear on his body and the several rules that they needed to follow

before the day of the experiment. The written consent was also collected from the participants at this time. They also got the information that they could quit from the experiment at any time. They would also answer three different questionnaires - Eysenck's personality questionnaire, Sniffin' Sticks, Odor awareness scale.

2. At the beginning of each of the 3 experiments these steps where always done: the participant would wash their underarm area with purified water; they would answer the POMS (Profile of Mood States), the VAS (100mm) measuring the type and intensity of the emotions experienced by the participants (anger, sadness, happiness, surprised, fear and disgust) and the STAI-State.

3. At the beginning of the first session of each session, the participant was asked to wash their hands with a blue soap (this was done because of the EDA sensor). The electrodes were placed on the body of the participant (the positioning was maintained through experiments). Signal quality was always visually checked.

4. After all the electrodes were positioned a time window of 10 minutes waiting time was considered in order to stabilise the signals.

5. All the triggers were deployed automatically via software, so, at the beginning of each segment the trigger would be deployed accordingly.

6. The user would then start the experiment. A first 5 minute video was shown, this was a baseline video that triggered little to no emotion to the user, as verified by the pilot experiment, this video served to record the baseline of the participant. This is when the baseline segmentation of the data is recorded.

7. After the conclusion of the first video some questions showed up to the user, by an analogue visual scale questionnaire. These questions were about how the participant felt during the video and how he thought the other participants felt watching the same video.

8. Concluding the questionnaire, the experiment would continue. On the first session a 20 minute movie was shown, this movie was rated on a pilot that didn't trigger strong emotions. This marked the Neutral and first part of the experiment. The neutral segmentation of the data is recorded here, this is one of the segments that will be used on the analysis and training phases.

9. With this done another questionnaire showed up, this was equal to the first one and asked the same set of questions.

10. Subsequently with the conclusion of the neutral part of the experiment the armpits of the participants would be cleaned up with the proper tools, so the second phase could start.

11. This is the randomised phase, it could start with the fear or the happiness, and the set of videos present on these phases were also at random positions.

12. The second phases would start again with a 5 minute baseline video followed by the same questionnaire as before.

13. Afterwards the emotional experiment would start, the participant watched a set of videos (concatenation of several movies inducing the same emotion). One of the emotions segments of the data is recorded here, this is another of the segments that will be used on the analysis and training phases.

14. Another questionnaire would show up and the participant would answer, and the first part of the experiment was concluded.

15. The second part of the experiment would start 2 hours apart from the ending of the first part of the study. In this part the electrodes were again positioned (maintaining the original position), and the participant would watch 5 minute baseline video, followed by emotionally set of videos. The other emotion segments of the data are recorded here, this is last segment that will be used on the analysis and training phases.

16. The ending of the second phase marked the ending of the experiment.

The videos for the happy and fear condition were selected from a pilot experiment. The neutral collection was only of a movie with a duration of 20 minutes.

With all the experiments done it was created a data set. This data set is constituted of two emotional components, being them Negative (Fear) and Positive (Happiness) and a combination of several physiological variables (ECG, Facial EMG, EDA and RSP).

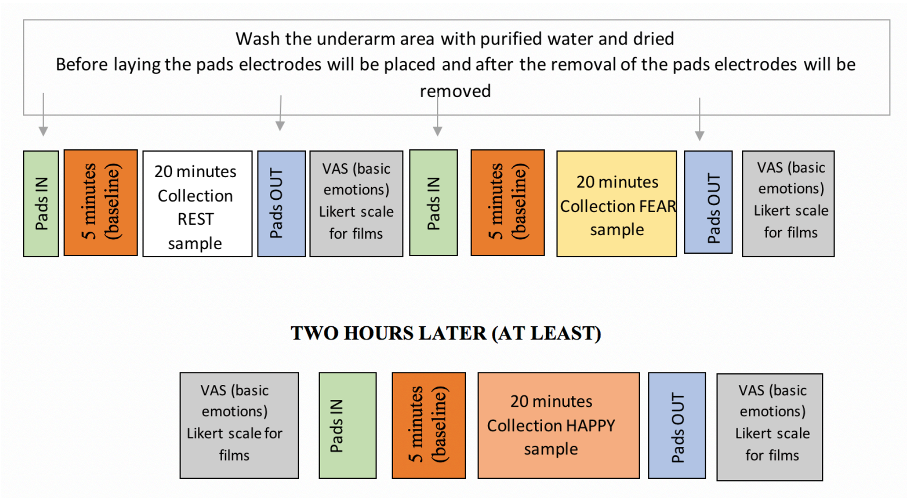


Figure 4.3: Experimental Protocol
(Source: Marta Rocha)

The data set provided was saved in ACQ files (created by the AcqKnowledge) and sampled at 200 Hz, which constitutes a problem for EMG analysis. To overcome this problem the EMG signal was resampled to 1000 Hz. Aware that 200 Hz for EMG data collection compromised some characteristics of the signal, it was decided to proceed with the analysis.

4.3 Analysis

As the data set was provided in ACQ files there was a necessity to convert them to .MAT, in order to be readable in python libraries.

The data processing was divided in several steps. The first one was the data preparation, by the segmentation of the usable data.

After having the files prepared with the necessary data it was create a python script that would read these files and put the data in the appropriate data models. All of the files were read, the information for each of the physiological signals was stored as well as the triggers and the sample they appeared. Next, based on the values present on the triggers, the previous signals were segmented, keeping the information for the baseline phase and the experimented phase, these two segments were then stored in dictionaries for each of the signals that had as key the id of the participant and the emotional experiment.

Filter

The data filtering was performed by the application of filters available on the *Neurokit* library [1] from python.

For resampling of the EMG signal, the method used was the Fourier method, by the application of the resample script from the scipy python library [2]. This method takes a signal with x samples and turns into a signal with xn samples, being the xn a number defined by the user, this transformation was based on the Fourier method.

[1] `https://neurokit.readthedocs.io`

[2] `https://docs.scipy.org/doc/scipy/reference/generated/scipy.signal.resample.html`

EMG

The EMG was filtered by means of a highpass butterworth filter of order 4 at 100 Hz [61].

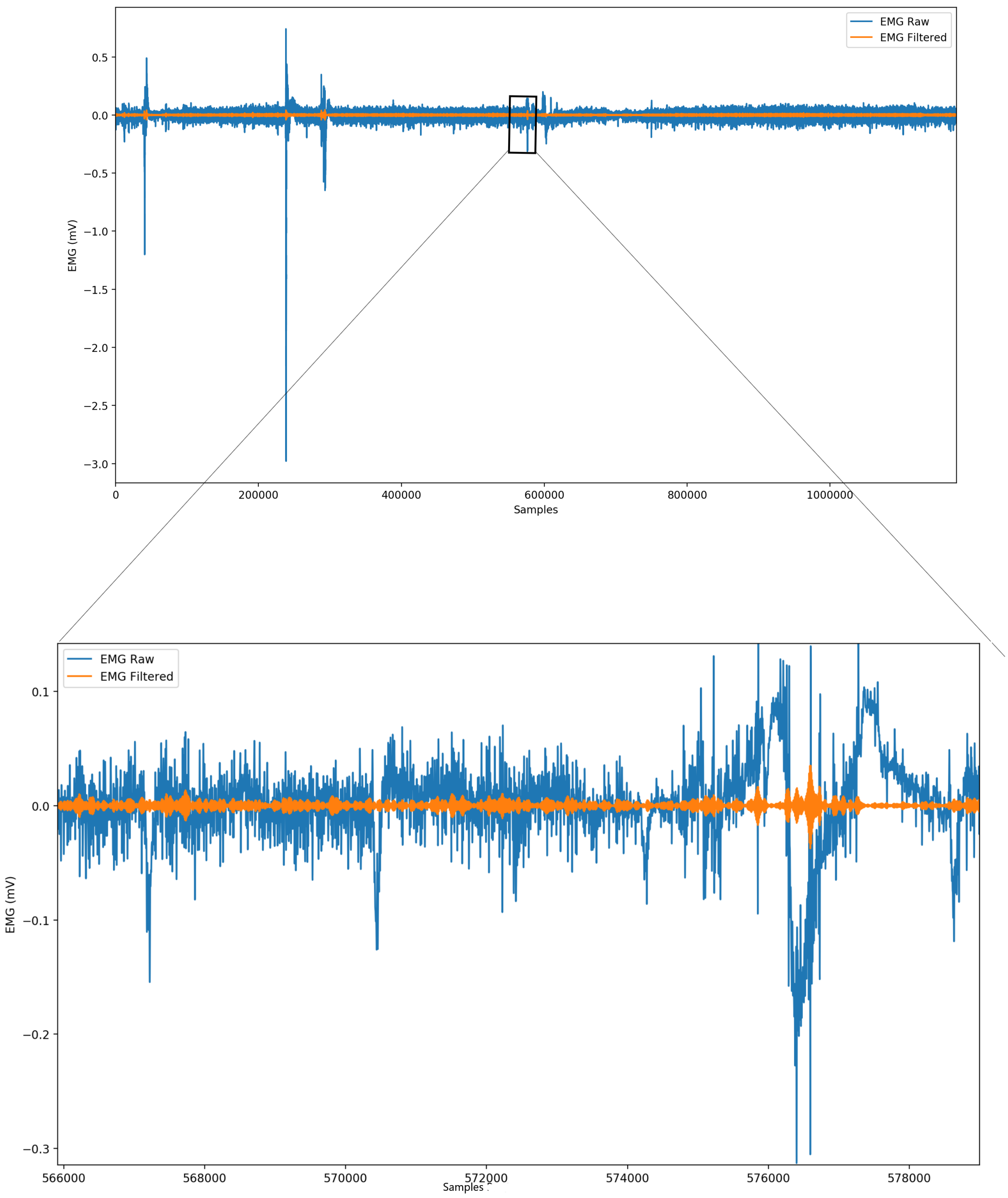


Figure 4.4: Comparison between Raw and Filtered EMG signals from participant 33

ECG

The ECG was filtered by means of a bandpass Finite Impulse Responde filter of order 66 (0.3 times the sampling rate) at 3Hz for the highpass and 45Hz for the lowpass [62].

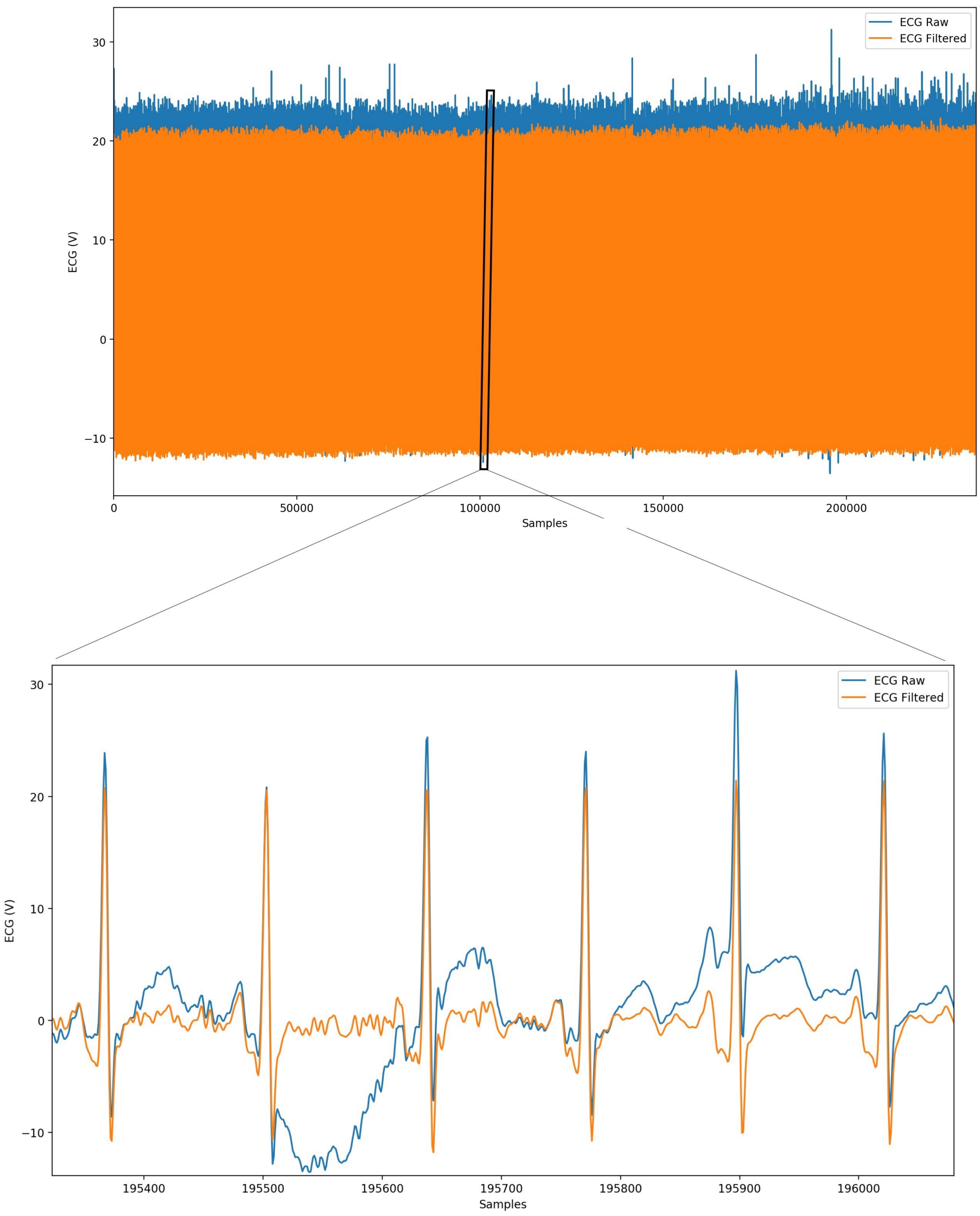


Figure 4.5: Comparison between Raw and Filtered ECG signals from participant 33

EDA
The EDA was filtered by means of a lowpass butterworth filter of order 4 at 5 Hz [63].

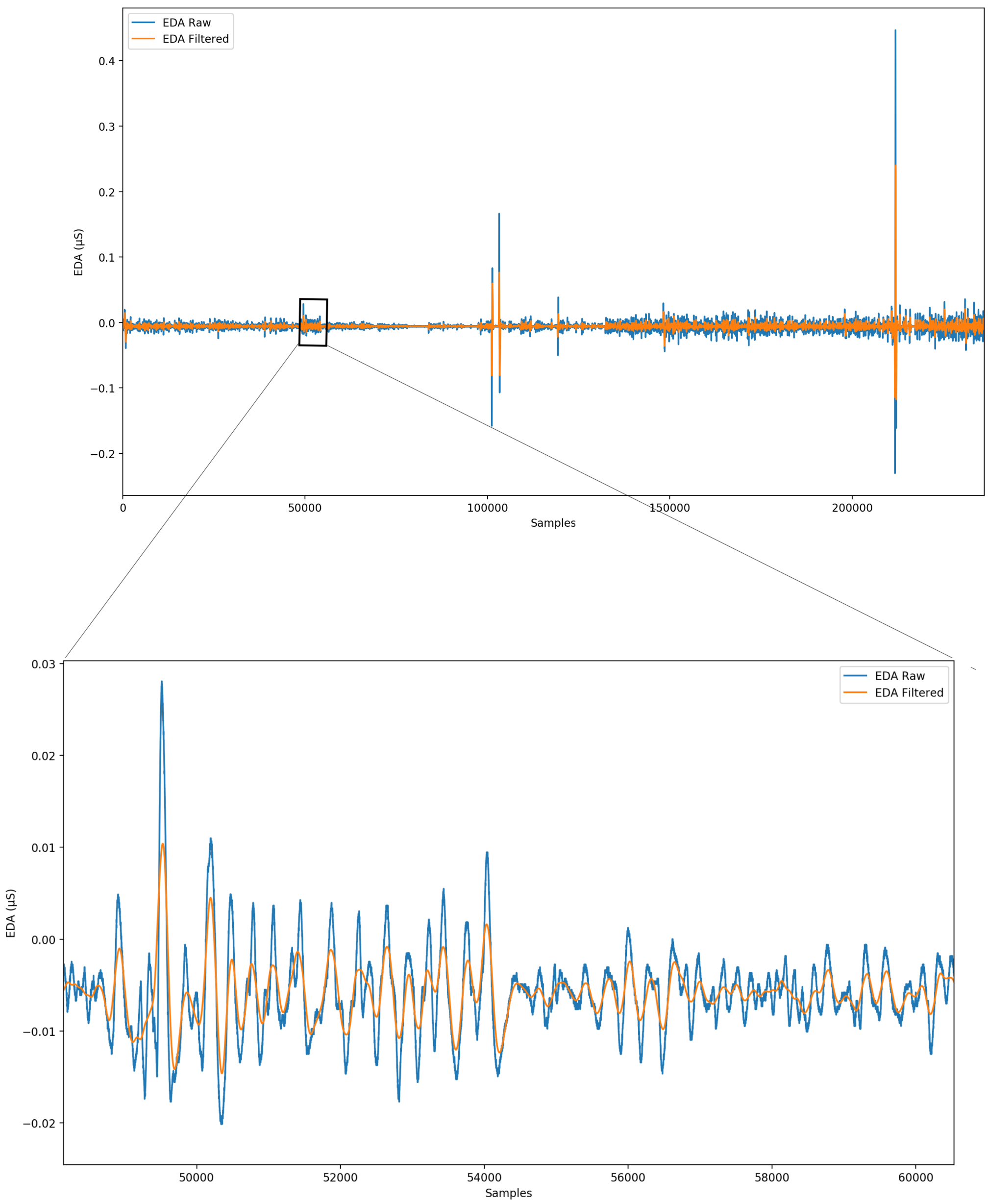


Figure 4.6: Comparison between Raw and Filtered EDA signals from participant 33

RSP

The RSP was filtered by means of a bandpass butterworth filter of order 2 at 0.1Hz for the highpass and 0.35Hz for the lowpass [64].

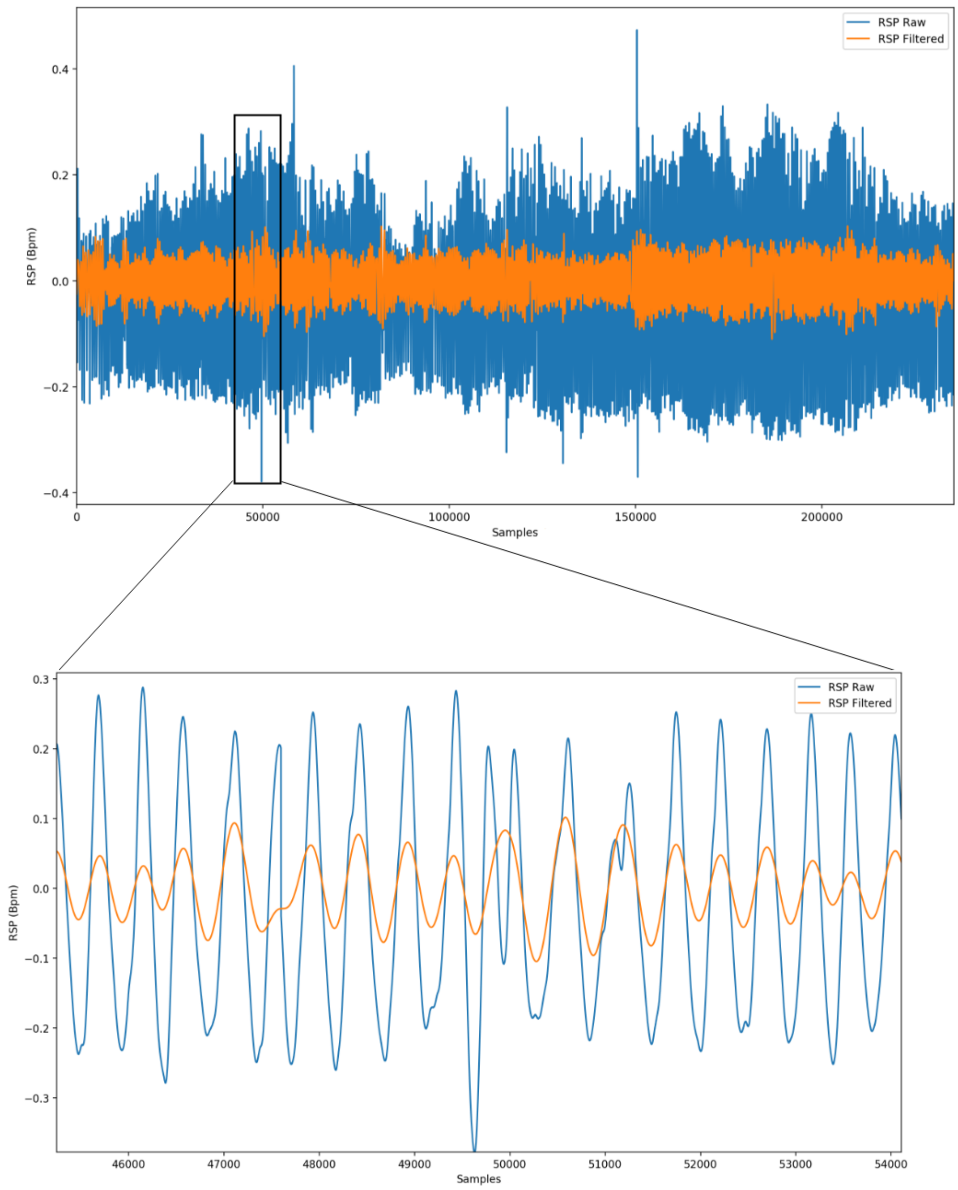


Figure 4.7: Comparison between Raw and Filtered RSP signals from participant 33

4.4 Emotion Feature Selection

The filtered data was analysed in order to extract the relevant features for emotion identification.The Neurokit [3] and Biopssy [4] libraries from Python were used in this procedure.

From each of the used signals several features were extracted (Table 4.4). After the extraction of all the features, it was necessary to study and understand, which of them are significant in emotion discrimination.

Variance Threshold

In this technique, the variance of each individual feature is calculated and only the ones that surpass a certain threshold are kept. The method idea is that in theory a feature with low variance, little information will give to the classifier. The method VarianceThreshold [5] from the *sklearn.feature_selection* library was used for this step.

Several values for threshold were tested between 0.9 and 0.05. The final value chosen was 0.1, the reason behind this decision was because this value gave the best trade off between the best number of features and quantity of information each one give.

The next tables (Table: 4.1) will show the selected features and their associated variance, while the table (Table: 4.4) shows the comparison between the features available in the library and the ones selected.

[3] `http://neurokit.readthedocs.io`

[4] `http://biosppy.readthedocs.io`

[5] `http://scikit-learn.org/stable/modules/generated/sklearn.feature_selection.VarianceThreshold.html`

Table 4.1: Signals Features Variance Threshold Selection

Signal	Signal Feature	Variance
ECG	Heart_Rate	1.1e01
	ECG_Systole	2.1e-01
	ECG_Signal_Quality	1.1e-01
	ECG_RR_Interval	2.6e09
	ECG_HRV_HF	2.1e08
	ECG_HRV_LF	2.6e08
	ECG_HRV_ULF	1.4e08
	ECG_HRV_VHF	7.9e07
	ECG_HRV_VLF	3.1e08
	R_Peaks	2.7e07
	Cardiac_Cycles	9.4e02
	T_Waves	2.7e07
	P_Waves	2.7e07
	Q_Waves	2.7e07
EDA	SCR_Onsets	2.9e09
	SCR_Peaks_Amplitudes	8.2e02
	EDA_Phasic	5.3e02
	EDA_Tonic	5.3e02
EMG	EMG_Pulse_Onsets	3.7e09
	EMG_Activation	2.1e-01
RSP	RSP_Rate	1.0e01
	RSP_Inspiration	2.5e-01
	Cycles_Onsets	4.5e09
	Expiration_Onsets	4.6e09

Correlation Threshold

This technique calculates the correlation matrix between all features of a given patient in a given condition. After the matrix is acquired all of the values are compared between them and the ones that are above a certain threshold are displayed to be selected. In theory if two or more features have a high correlation, means that the information is redundant, so one of them can be discarded.

The values used for the threshold selection were tested between 0.9 and 0.05, this values were chosen based on the number of features they output (the middle values little difference made). The final value chosen was 0.1 (table 4.2), the reason behind this decision was because this value gave the best combination between number of features and correlation.

Table 4.2: Signals Correlated features

Signal	Correlated features	Not selected feature
ECGA	ECG_HRV_HF, ECG_HRV_LF, ECG_HRV_VLF, Heart_Rate	ECG_HRV_VLF
	ECG_HRV_ULF, ECG_RR_Interval	ECG_HRV_ULF
	Cardiac_Cycles, T_Waves, P_Waves, Q_Waves, R_Peaks	Cardiac_Cycles
EDA	SCR_Peaks_Indexes, SCR_Recovery_Indexes, SCR_Onsets	SCR_Recovery_Indexes
EMG	EMG_Activation, EMG_Envelope	EMG_Envelope
RSP	Expiration_Onsets, Cycles_Onsets	Cycles_Onsets

PCA - Principal Component Analysis

This technique goal is feature reduction. Therefore, it combines the features to preserve the maximum information with a lower number of components. Basically, the method applies a linear dimensional reduction using Singular Value Decomposition of the data to project it to a lower dimensional space. The Principal Component Analysis (PCA) method from the sklearn library [6] was used, as it is very commonly used both on data analysis and on the large number of articles read. The features that passed through the Variance Threshold and Correlation threshold are the ones passed to this method and based on the results present on the table 4.3, it was chosen 4 principal components and an explained variance of 95%, as it was the most preserver.

Table 4.3: Results for components that give an explained variance higher than 95%

Number of Components	Percentage of Components that together have variance higher than 95%
1	1.82%
2	9.26%
3	74.07%
4	14.81%

The table 4.4 presents the final selected features before the PCA that are going to be used as input in the machine learning sub block.

[6] http://scikit-learn.org/stable/modules/generated/sklearn.decomposition.PCA.html

Table 4.4: Selected Features

	All Features	Features Selected
EDA	SCR_Onsets, SCR_Peaks_Indexes, SCR_Recovery_Indexes, SCR_Peaks_Amplitudes, EDA_Raw, EDA_Filtered, EDA_Phasic,EDA_Tonic, SCR_Recoveries,SCR_Peaks	SCR_Onsets, EDA_Phasic, EDA_Tonic, SCR_Peaks_Amplitudes
EMG	RSP_Filtered, RSP_Rate, RSP_Inspiration, Expiration_Onsets, Cycles_Length	EMG_Activation, EMG_Pulse_Onsets
RSP	RSP_Raw, RSP_Filtered, RSP_Rate, RSP_Inspiration, Cycles_Onsets, Expiration_Onsets, Cycles_Length, Respiratory_Variability	RSP_Rate, RSP_Inspiration, Expiration_Onsets, Cycles_Length
ECG	R_Peaks, Cardiac_Cycles, T_Waves, P_Waves, Q_Waves, Cardiac_Cycles_Signal_Quality, Average_Signal_Quality, ECG_Signal_Quality, ECG_Raw, ECG_Filtered, ECG_R_Peaks, Heart_Rate, ECG_Systole, ECG_Signal_Quality, ECG_RR_Interval, ECG_HRV_HF, ECG_HRV_LF, ECG_HRV_ULF, ECG_HRV_VHF, ECG_HRV_VLF, n_Artifacts, RR_Intervals, RMSSD, meanNN, sdNN, cvNN, CVSD, medianNN, madNN, mcvNN, pNN50, pNN20, Triang, Shannon_h, ULF, VLF, LF, HF, VHF, Total_Power, LFn, HFn, LF/HF, LF/P, HF/P, DFA_1, DFA_2, Shannon, Sample_Entropy, Correlation_Dimension, Entropy_Multiscale_AUC, Entropy_SVD, Entropy_Spectral_VLF, Entropy_Spectral_LF, Entropy_Spectral_HF, Fisher_Info, FD_Petrosian, FD_Higushi	R_Peaks, T_Waves, P_Waves, Q_Waves, ECG_Signal_Quality, Heart_Rate, ECG_Systole, ECG_RR_Interval, ECG_HRV_HF, ECG_HRV_LF, ECG_HRV_ULF, ECG_HRV_VHF, ECG_HRV_VLF

4.5 Emotion Classification

After feature identification, the next step was the classification part.

As the EMG is sampled at 200 Hz it can cause problems as it is a very low frequency for a signal that as an optimal recording frequency of 800-1000 Hz, as it was proved that sampling the EMG at 500 and 250 Hz resulted in statistically significant effects on onset latency and burst duration measures, it also results in decrease of the peak amplitude, concluding that with this the waveform shape may change to such an extent that the appearance and timing of the peaks is dramatically altered. [65].

For classification purpose the original dataset was divided in different ways between train and test data groups. The First one in the training consider cutting the full experiment in 1 minute cuts with PCA after the feature selection. The second one is also only considering 1 minute data in training, the difference was that the classification was performed before PCA was applied. The third and fourth one are equal to the first and second but in this case no EMG was used as the previous realisation of the sampling rate on the EMG made it not viable.

The methods used for the classification part were: K-nearest neighbours, Neural Network, this were the chosen based on some papers used as inspiration.

To counterbalance the test and train split, the same participant is never in both of them. All the conditions of the user are putted on the train or on the test split and never in both.

For the split of the data 15% of the participants were selected for the test.

The method performance was inferred by:

- Sensitivity - evaluates the probability of the method classifying an emotion in a class, when it effectively belongs to it;
- Specificity - evaluates the probability of the method not classifying an emotion in a specific class, when it does not belong to it;
- Accuracy - evaluates the closeness of the predicted emotions vs the actual value of the emotion;
- Confusion Matrix - is a table that is often used to describe the performance of a classifier, it allows easy identification of confusion between classes, that is, the common mislabel of one class as the other. It creates a summary of prediction results on a classification problem;

Tuning the parameters

The classifier methods parameters were tuned using GridSearchCV method from the *scikit-learn* library [7]. This method is the most traditional to perform parameter optimisation. The main idea is to do an exhaustive search through a manually specified set of

[7] `http://scikit-learn.org/stable/modules/generated/sklearn.model_selection.GridSearchCV.html`

parameters of a learning algorithm. At the end it will give the performance of each of the fit of the models with the different parameters and the optimal can be chosen.

Neural Network

Artificial Neural Networks are algorithms that are manly based on the human brain neurons and synapses. The main goal of this algorithms is to recognise patterns. As they receive an input they will try to label or cluster it. A neural network is constituted of several layers (always one input layer, that has the size of the feature vector, one output layer that will have the size equal to the number of classes, and zero or more hidden layers), each layer is constituted of several units called neurons (these neurons hold a value inside of it called "activation"). These neurons are connected to the neurons on the next layer, by weighted connections, these connections receive a random weight (value). This weight will be multiplied by the value of the activation of the neuron that is connected to it and will then give an activation value to the neuron in the next layer (each of the neurons of the next layer is connected to all of the neurons in the previous layer). These weights are always random, and the main goal of the network is to find the best values that lead to the desired output.

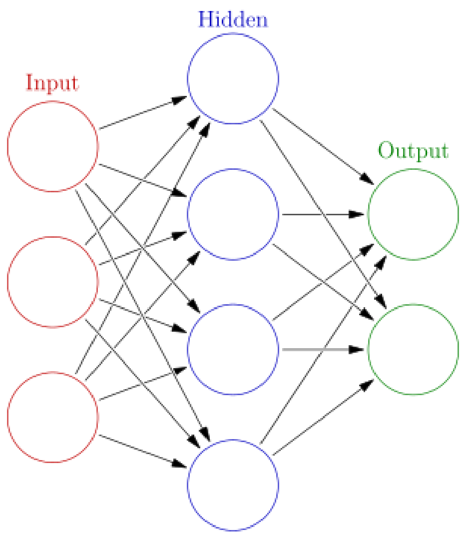


Figure 4.8: Articial Neural Network
(Source: https://goo.gl/hLJT7i)

The Neural Network was created using the Keras library [8] from python. This library provides several back ends (TensorFlow, Theano, CNTK), the one used is the TensorFlow, as it is the most popular on the internet and in some of the articles.

The model created to classify the given data set is the described in fig. 4.9, as outputted by the library method.

[8] `https://keras.io`

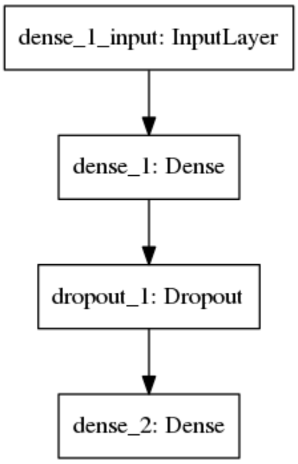


Figure 4.9: Neural Network structure for one minute division experience with no PCA

Parameter optimization

The gridsearch started from a set of values for each of the parameters, that were previously manually selected, the network would run with each combination of values, and, at the end, the parameters for the network with the best error were the selected ones, the selected parameters are present on the following enumeration:

Number of nodes / Number of hidden layers The utility of this parameter is to construct the structure of the network, the number of layers will characterise how deep is the network / how many levels needs the input pass through, so it gets the output value. The number of nodes / neurons is the depth / number of levels of each individual hidden layer.

Number of nodes per layer for the final topology as selected by the gridsearch:

- Input layer - 35 nodes for the 1 minute with PCA, 15 for the 1 minute without PCA, 50 nodes for the 1 minute with PCA without EMG and 20 for the 1 minute without PCA without EMG.
- Output layer - 3 nodes (3 classes)

Optimizer

Optimisation algorithms are algorithms that help the neural network to minimise the Error function of the prediction. There were several algorithms with similar results. So, the one selected by the Grid Search was the one chosen, *Adam* [9].

Activation

Each node has an activation function, where it is defined the output of that node give a certain input or set of them.

[9] `https://www.tensorflow.org/api_docs/python/tf/keras/optimizers/Adam`

Activation function per layer:

- Input layer - *sigmoid*, Rectified Linear Unit, is the more common function, the objective of this function is to return 0 if the value is negative and return the value if otherwise [66].
- Output layer - *softmax*, Normalized exponential function, it is a generalisation of the logistic function that takes a K-dimensional vector $\mathbf{x}$ of random real values and transforms it to a K-dimensional vector $\sigma(\mathbf{x})$ of real values, where each value is between the range of (0,1) and all the values add up to 1 [67].

Train and test

With all the parameters optimised it was verified if the model was over fitting the data, to these 100 runs were made with the best parameters.

For the 1 minute cut with PCA a mean error of 0.651 $\pm$ 0.023 was achieved. For the 1 minute cut without PCA a mean error of 0.666 $\pm$ 0.004 was achieved. For the 1 minute cut with PCA without EMG a mean error of 0.663 $\pm$ 0.015 was achieved. For the 1 minute cut without PCA and without EMG a mean error of 0.598 $\pm$ 0.039 was achieved. As the standard deviation of all of the models was low it can be concluded that there wasn't any kind of overfitting.

To train the model it was used 85% of all of the data and the rest was used to test 50 epochs were used to run the network and try to stabilise it, or when the loss of the validation was the same for 3 runs.

The best model was stored, and his performance was checked in the 1 minute cut with PCA got an overall accuracy of 0.446. The 1 minute cut without PCA got an overall accuracy of 0.453. The 1 minute cut with PCA without EMG got an overall accuracy of 0.391. The 1 minute cut without PCA and without EMG got an overall accuracy of 0.509. The metrics at the table 4.5 for the 1 minute cut with PCA, at 4.7 for the 1 minute cut without PCA, at the table 4.9 for the 1 minute cut with PCA without EMG and at 4.11 for the 1 minute cut without PCA without EMG. This performance is discussed in the next few lines and the confusion matrix can be checked at table 4.6 for the 1 minute cut with PCA, at table 4.8 for the 1 minute cut without PCA, at table 4.10 for the 1 minute cut with PCA without EMG and at table 4.12 for the 1 minute cut without PCA without EMG.

Table 4.5: NN results of the best 1 minute cut experiment with nn with PCA

	Sensitivity	Specificity
Neutral	0.842	0.278
Fear	0.0	1.0
Happy	0.495	0.889

Table 4.6: Confusion Matrix for the best 1 minute cut experiment with nn with PCA

Real \Predicted	0 - Neutral	1 - Fear	2 - Happiness
0 - Neutral	80	0	15
1 - Fear	89	0	6
2 - Happiness	48	0	47

Table 4.7: NN results of the best 1 minute cut experiment with nn without PCA

	Sensitivity	Specificity
Neutral	0.337	0.751
Fear	0.425	0.621
Happy	0.411	0.712

Table 4.8: Confusion Matrix for the best 1 minute cut experiment with nn without PCA

Real \Predicted	0 - Neutral	1 - Fear	2 - Happiness
0 - Neutral	32	31	32
1 - Fear	19	48	28
2 - Happiness	15	41	39

Table 4.9: NN results of the best 1 minute cut experiment with nn with PCA without EMG

	Sensitivity	Specificity
Neutral	0.495	0.611
Fear	0.358	0.695
Happy	0.326	0.784

Table 4.10: Confusion Matrix for the best 1 minute cut experiment with nn with PCA without EMG

Real \Predicted	0 - Neutral	1 - Fear	2 - Happiness
0 - Neutral	47	32	16
1 - Fear	36	34	25
2 - Happiness	38	26	31

Table 4.11: NN results of the best 1 minute cut experiment with nn without PCA without EMG

	Sensitivity	Specificity
Neutral	0.821	0.463
Fear	0.179	0.937
Happy	0.526	0.863

Table 4.12: Confusion Matrix for the best 1 minute cut experiment with nn without PCA without EMG

Real \Predicted	0 - Neutral	1 - Fear	2 - Happiness
0 - Neutral	78	7	10
1 - Fear	62	17	16
2 - Happiness	40	5	50

K-nearest neighbours

k-Nearest Neighbours is one of the simplest machine learning algorithms, the main objective of this algorithm is to categorise an input. To do this task it uses the k nearest neighbours, that is, when a new point arrives at the algorithm this will look to the k nearest neighbours and based on the classification of those nearest neighbours it will classify the new point.

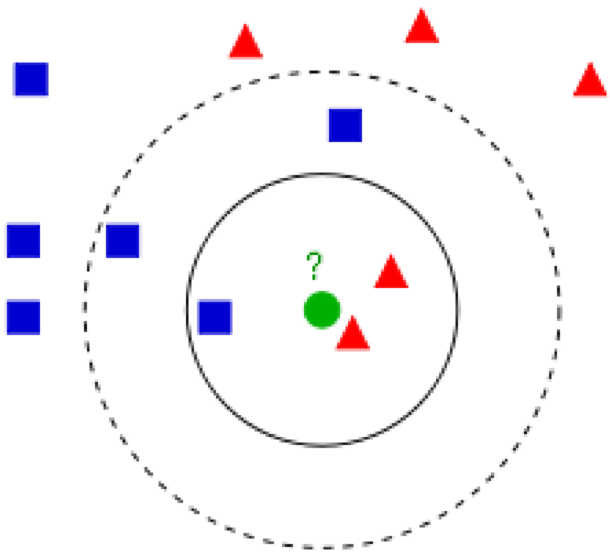


Figure 4.10: k-Nearest Neighbour
(Source: https://goo.gl/n7nivf)

To create this model the KNeighborsClassifier from the *sklearn* library was used.

These methods requires two parameters, number of neighbours and weight function. To choose the best for the data being used several tests were done and are presented in the next point.

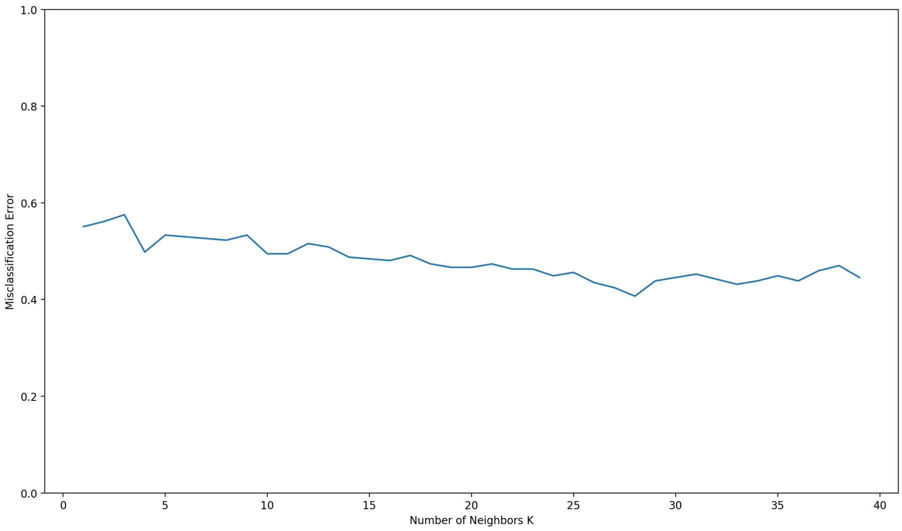
Parameter optimisation

n_neighbours

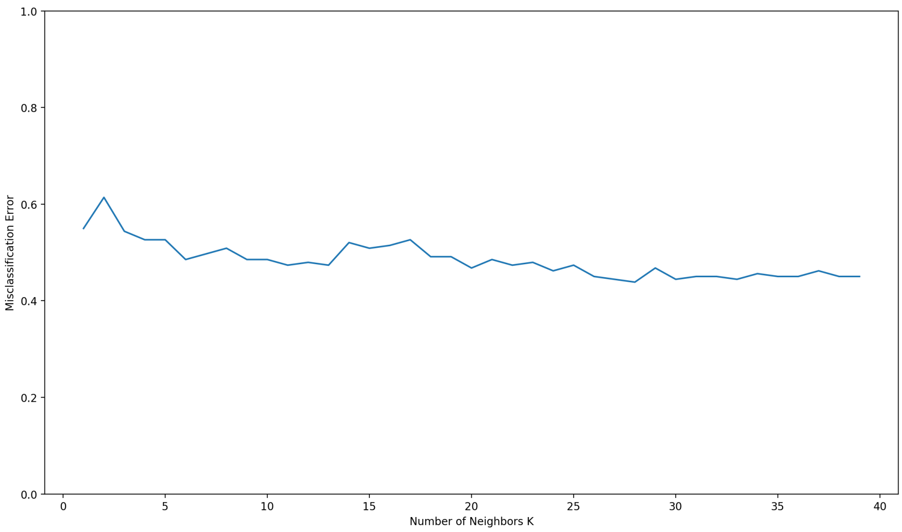
This parameter represents the number of neighbours the point should be close to, to belong to that neighbour's class.

To choose this parameter all the values neighbours values between 1 and 40 were tested and the one with the best error were selected, the results are shown on fig. 4.11a, fig. 4.11b, fig. 4.11c and fig. 4.11d.

For 1 minute cut with PCA data set it was selected k equals to 28, as it was the one that best performed as it can be checked on fig. 4.11a, for 1 minute cut without PCA data set it was also selected the k equal to 28, as it can be checked on fig. 4.11b, for 1 minute cut with PCA data set without EMG it was selected the k equal to 27, as it can be checked on fig. 4.11c and for 1 minute cut without PCA data set without EMG it was selected the k equal to 28, as it can be checked on fig. 4.11d. These were the selected values but as it can be checked on the figures all of them are very stable.

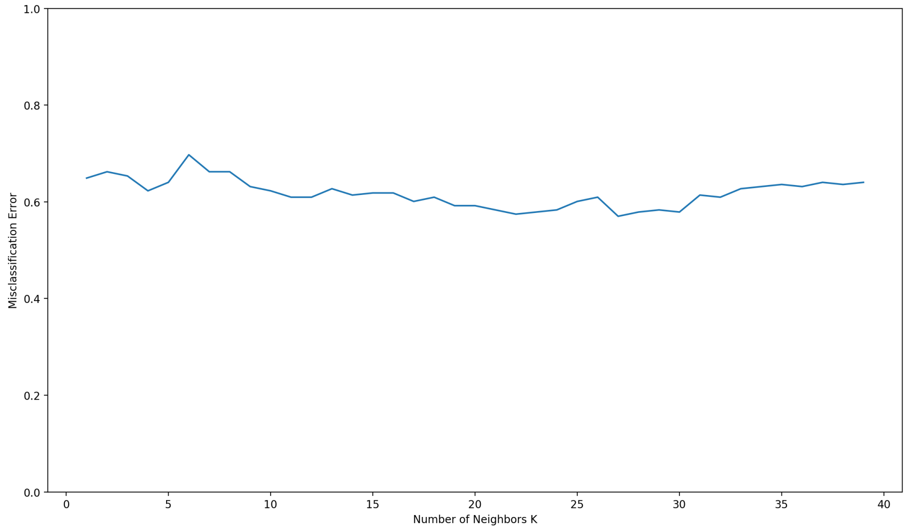


(a) Number of neighbours and associated error for the 1 minute cut data

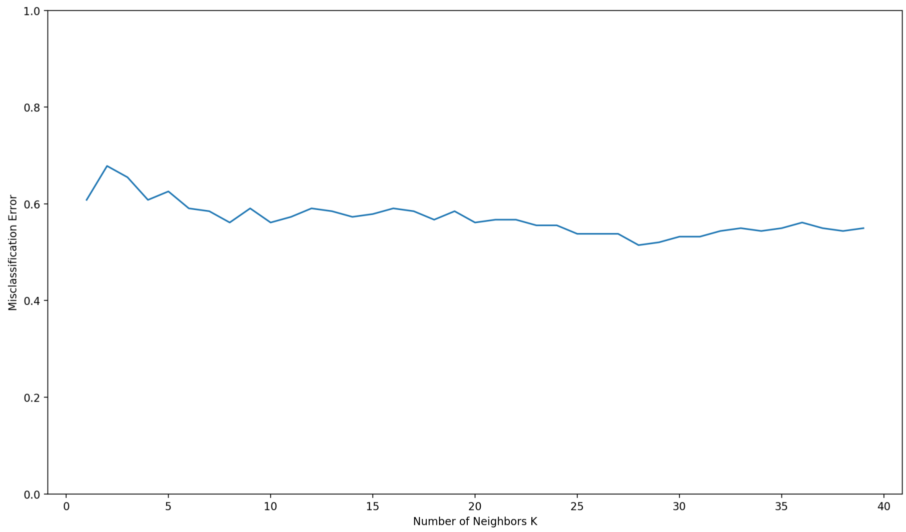


(b) Number of neighbours and associated error for the 1 minute cut data withuout PCA

Figure 4.11: Number of neighbours and associated error



(c) Number of neighbours and associated error for the 1 minute cut data with PCA without EMG



(d) Number of neighbours and associated error for the 1 minute cut data withouut PCA without EMG

Figure 4.11: Number of neighbours and associated error continued

weight

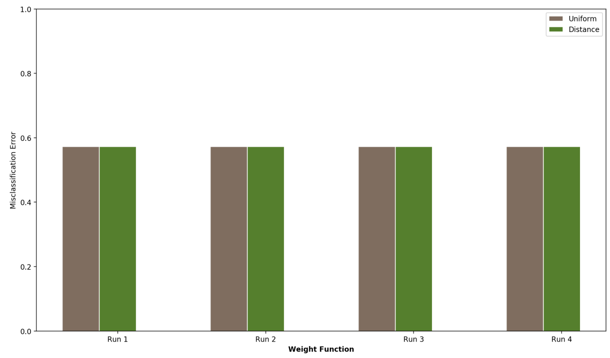
This parameter is to select the weight function to be used in prediction.

There were two possible functions to choose from:

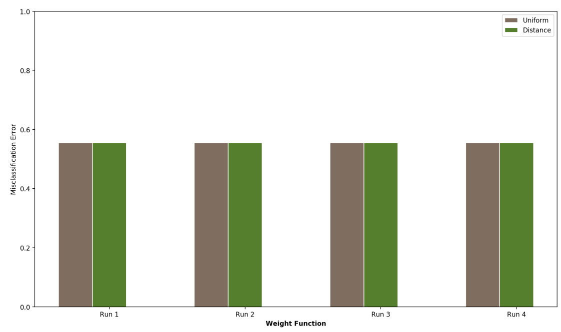
- Uniform: Uniform weights. All points in each neighbourhood are weighted equally.
- Distance: Weight points by the inverse of their distance. In this case, closer neighbours of a query point will have a greater influence than neighbours which are further away.

To choose this parameter each function was tested 4 times for each of the data sets.

The results are shown on the next graph 4.12a, 4.12b, 4.12c and 4.12d. As all performed the same as it can be checked on the figures, the decided one was the default one, being it the Uniform, since it is the simplest one.

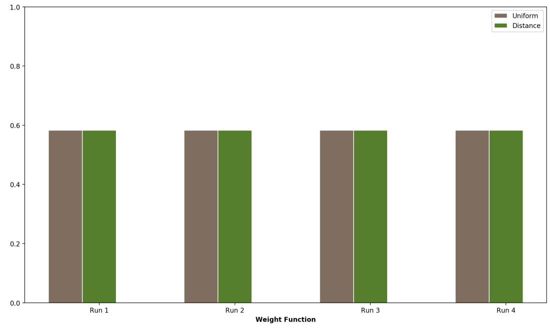


(a) Weight function and associated error for the 1 minute cut data

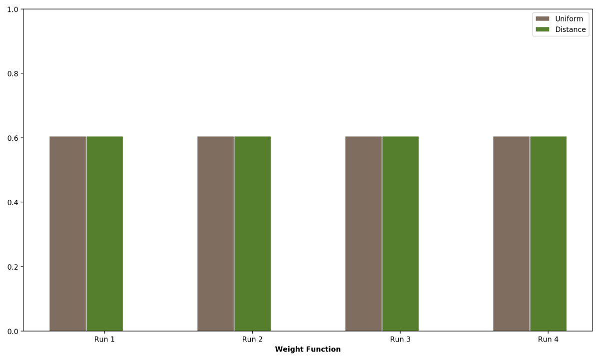


(b) Weight function and associated error for the 1 minute cut data without PCA

Figure 4.12: Weight function and associated error



(c) Weight function and associated error for the 1 minute cut data without EMG



(d) Weight function and associated error for the 1 minute cut data without PCA without EMG

Figure 4.12: Weight function and associated error continued

Train and test

With all the parameters optimised it was verified if the model was over fitting the data, to do these 100 runs were done and the mean error and standard deviation were verified. For the 1 minute cut with PCA a mean error of 0.627 $\pm$ 1.110e-16 was achieved. For the 1 minute cut without PCA a mean error of 0.578 $\pm$ 2.220e-16 was achieved. For the 1 minute cut with PCA and without EMG a mean error of 0.612 $\pm$ 0.028 was achieved. For the 1 minute cut without PCA and without EMG a mean error of 0.569 $\pm$ 0.092 was achieved. As the standard deviation of all of the models was low it can be concluded that there wasn't any kind of overfitting.

To train the model it was used 85% of all the data and the rest was used to test.

The best model was stored, and his performance was checked in the 1 minute cut with PCA got an overall accuracy of 0.486. The 1 minute cut without PCA got an overall accuracy of 0.502. The 1 minute cut with PCA without EMG got an overall accuracy of 0.498. The 1 minute cut without PCA without EMG got an overall accuracy of 0.614. The metrics at the table 4.13 for the 1 minute cut with PCA, at table 4.15 for the 1 minute cut without PCA, at table 4.17 for the 1 minute cut with PCA without EMG and at table 4.19 for 1 minute cut without PCA without EMG. This performance is discussed in the next few lines and the confusion matrix can be checked at table 4.14 for the 1 minute cut with PCA, at table 4.16 for the 1 minute cut without PCA, at table 4.18 for the 1 minute cut with PCA without EMG, at table 4.20 for the 1 minute cut without PCA without EMG.

Table 4.13: Results of the best run for the full experiment with knn with PCA

	Sensitivity	Specificity
Neutral	0.487	0.717
Fear	0.447	0.704
Happy	0.526	0.809

Table 4.14: Confusion Matrix for the best full experiment with knn with PCA

Real \Predicted	0 - Neutral	1 - Fear	2 - Happiness
0 - Neutral	37	23	16
1 - Fear	29	34	13
2 - Happiness	14	22	40

Table 4.15: Results of the best run for the full experiment with knn with no PCA

	Sensitivity	Specificity
Neutral	0.579	0.647
Fear	0.516	0.647
Happy	0.411	0.958

Table 4.16: Confusion Matrix for the best full experiment with knn with no PCA

Real \Predicted	0 - Neutral	1 - Fear	2 - Happiness
0 - Neutral	55	38	2
1 - Fear	40	49	6
2 - Happiness	27	29	39

Table 4.17: Results of the best run for the full experiment with knn with PCA without EMG

	Sensitivity	Specificity
Neutral	0.663	0.674
Fear	0.432	0.753
Happy	0.4	0.821

Table 4.18: Confusion Matrix for the best full experiment with knn with PCA without EMG

Real \Predicted	0 - Neutral	1 - Fear	2 - Happiness
0 - Neutral	63	23	9
1 - Fear	29	41	25
2 - Happiness	33	24	38

Table 4.19: Results of the best run for the full experiment with knn with no PCA without EMG

	Sensitivity	Specificity
Neutral	0.674	0.716
Fear	0.516	0.837
Happy	0.653	0.868

Table 4.20: Confusion Matrix for the best full experiment with knn with no PCA without EMG

Real \Predicted	0 - Neutral	1 - Fear	2 - Happiness
0 - Neutral	64	15	16
1 - Fear	37	49	9
2 - Happiness	17	16	62

4.6 Results summary

Analysing the results, the following remarks can be noted:

- As the standard deviation of the error on the 100 runs is low it can be said that the algorithm is not over fitting the data;

- The fear emotion when no EMG is present is the emotion most difficult to classify;

- The knn model was the one with the best results with any of the data sets;

- The data without PCA in both methods displayed the best results, as 4 principals components showed up to be not sufficient, so the machine learning could understand the data;

- The data set without EMG got the best results has the sampling rate was not the optimal;

With the Knn algorithm the 1 minute cut without PCA was the method that showed best results overall.

The results for the other classifiers never got the 50% mark of accuracy, even after the grid search was done no great results were found. Possibly the main reason that caused these results was the data set, due to the lowest sampling frequency. Some of the ideas for the future work to improve the data would be:

- Divide the data in different time frames, as the 1 minute time frame has been concluded to not give the best results, so, a division into 2 or 4 minutes would probably improve the accuracy values of the classifiers;

- Trying different models from the classification section, or even try a few from other sections. With this probably some improvements would be visible with more robust algorithms;

- Selecting different features from each of the different signals, or even reducing the number of the ones that were selected to be more precise and all with more information;

- Create features by hand as the method in the library is rigid and it is possible to find features that best represent the psychophysiology of emotion.

- Explore parameter free data mining techniques;

Chapter 5

Conclusions

The system showed promising results when parts of these solutions were integrated, as it can be seen through the work, an emotion identification system is possible and viable, and with the proper sensor it can be used in a daily life basis. As this is an early proof of concept in a bigger project some problems appeared. The application of such system can make a difference especially for people with special needs.

All the 4 challenges present on section 1.2 were answered, a full end-to-end system was created, since the monitoring phase, until the feedback phase passing by the feature extraction and machine learning phases. A non-intrusive sensor was evaluated (Vital Jacket), and it was proved as usable in real life daily routine. The relation between physiological signals and emotion was understood. Based on the feature selection some relevant physiological features were selected, being used as input in machine learning models to classify emotions.

The low performance present on both classification methods could be caused by several issues that are going to be described, also some possible solutions for this problem are going to be presented. The major techniques used for data exploration and feature selection/reduction are the most standard/classic used in most of the articles read, the features used for each of the signals were also manly the ones used on most of the articles, give this, a bigger exploration of the signals could produce different results. Also, using a different set of features can produce different results A more in-depth exploration of the data is needed. The similarity between emotions may be explored by means of different metrics (such as parameter free data mining). Also, the study of relations between emotions, considering different persons is of valuable exploration, since it will allow to interpret possible relations and help to describe the psychophysiological pattern of emotions.

The sampling rate of the database is 200 Hz, which can also cause problems in the signal processing, especially for the EMG signal. 200Hz is a very low frequency for a signal that has an optimal recording frequency of 800-1000 Hz [65]. Sampling the EMG at 500 and 250 Hz resulted in statistically significant effects on onset latency and burst duration measures, it also results in decrease of the peak amplitude [65]. As for the ECG counterpart the recommended sampling frequency is of 250- 500 Hz or higher for HRV measurements without interpolation. As poor-quality records due to a low sampling rate can produce inaccurate measures and induce an error in the HRV spectrum, lower sampling rate may behave satisfactorily only if an algorithm of interpolation is used to refine the R wave [68]. So, using a different database with a higher sampling rate can improve in some way the model's performance.

Other identification algorithms should be tested, such as SVM, in order to validate if they produce a different algorithm performance.

Other problem that was caused by the data provided was the size of it, being that there were only 44 files for each emotion, and some of them had problems caused by the sensors. Therefore, only 37 files per emotion was usable. So, when splitting the data between train and test only around 30 files were used to train the methods, which may cause the system not to acquire the necessary information to identify patterns in emotion.

The ECG signal from the Vital Jacket presents a random delay that cannot be controlled. With this a full synchronisation can never be done in real time. However, when small windows are analysed the signals are comparable. Also, analysing the Heart Rate (HR), the results present almost 100% viability when compared to the Biopac HR.

The lower values acquired during the comparison phase can also be due to the difference in leads configuration used on the two systems, which caused a significant difference between the ECG Waves.

On fear condition, participants were reported of putting their hands in front of the face. This movements can cause noise on the ECG signal that filtering probably couldn't compensate for. Concluding that there were not any major problems, the heart rate from both sensors can be compared, and the Vital Jacket can be used in a daily life basis without the loose of its veracity, the ECG with the necessary synchronisation tools can also, be compared.

Some improvements should be done, and other data mining techniques should be planned, to improve the classification results. Nevertheless, the present system is able to collect, process and analyse physiological data in order to identify an emotion, as an end-to-end solution.

5.1 Future Work

As future work some ideas and thoughts are:

- Add more sensors to the android app, preferably wearable ones. The idea is to have an aggregation module that collect and organize the data. With this implementation the system will be less intrusive, which is a requisite for children with ASD;
- Validate the system with children with ASD. This will be important as the system was developed with data from young adults with normal development, this can result in inaccurate results when using it in ASD children;
- Create and train more robust machine learning algorithms;
- Use different feature extraction techniques;

- Validate the system performance on a different data set;

References

[1] "This is how your daily routine should look." [Online]. Available: https://www.inc.com/business-insider/healthy-daily-routine-habits-according-to-science.html

[2] "The world health report 2002 - reducing risks, promoting healthy life." [Online]. Available: http://www.who.int/whr/2002/en/

[3] F. D'Acquisto, "Affective immunology: where emotions and the immune response converge," *Dialogues in clinical neuroscience*, vol. 19, no. 1, p. 9, 2017.

[4] "How anxiety toys with your emotions." [Online]. Available: https://www.calmclinic.com/anxiety/anxiety-emotions

[5] "Human centred pervasive systems." [Online]. Available: http://inforum.org.pt/INForum2017/sessoes/human-centred-pervasive-systems.html

[6] M. Soegaard and R. Friis Dam, *The Encyclopedia of Human-Computer Interaction, 2nd Ed.*, 01 2013.

[7] "Emotions and types of emotional responses." [Online]. Available: https://www.verywellmind.com/what-are-emotions-2795178

[8] "Sensors." [Online]. Available: https://www.slideshare.net/RockHealth/rock-report-sensors-9962927/31-get_in_touchresearchrockhealthorgrockhealth

[9] "The role of sensors in our daily lives." [Online]. Available: http://archive2.cra.org/ccc/events/past-events/311-the-role-of-sensors-in-our-daily-lives

[10] "How big data is empowering ai and machine learning?" [Online]. Available: https://hackernoon.com/how-big-data-is-empowering-ai-and-machine-learning-4e93a1004c8f

[11] G. Blandford, A. Fouillé, and W. James, *What is an Emotion?: The Physiology of Emotion.* Editions Le Mono, 2017. [Online]. Available: https://books.google.pt/books?id=p6QqDwAAQBAJ

[12] K. Cherry, "Overview of the 6 major theories of emotion." [Online]. Available: https://www.verywellmind.com/theories-of-emotion-2795717

[13] C. Darwin, P. Ekman, and P. Prodger, *The Expression of the Emotions in Man and Animals.* Oxford University Press, 1998. [Online]. Available: https://books.google.pt/books?id=TFRtLZSHMcYC

[14] C. L. Lisetti and F. Nasoz, "Using noninvasive wearable computers to recognize human emotions from physiological signals," *EURASIP journal on applied signal processing*, vol. 2004, pp. 1672–1687, 2004.

[15] J. Benedek and R. Hazlett, "Incorporating facial emg emotion measures as feedback in the software design process," 01 2005.

[16] P. Ekman, R. Levenson, and W. Friesen, *Autonomic Nervous System Activity Distinguishes Among Emotions.* American Association for the Advancement of Science, 1983. [Online]. Available: https://books.google.pt/books?id=sMvynQEACAAJ

[17] J. T. Lanzetta and S. Orr, "Excitatory strength of expressive faces. effects of happy and fear expressions and context on the extinction of a conditioned fear response," vol. 50, pp. 190–4, 01 1986.

[18] V. S. R., C. B. N., and L. P. J., "Fear imagery and text processing," *Psychophysiology*, vol. 23, no. 3, pp. 247–253. [Online]. Available: https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1469-8986.1986.tb00626.x

[19] C. Smith and P. C. Ellsworth, "Patterns of cognitive appraisal in emotion," vol. 48, pp. 813–38, 05 1985.

[20] S. Vrana, "The psychophysiology of disgust: Differentiating negative emotional contexts with facial emg," vol. 30, pp. 279–86, 06 1993.

[21] R. Sinha, "Multivariate response patterning of fear and anger," *Cognition and Emotion*, vol. 10, no. 2, pp. 173–198, 1996. [Online]. Available: https://doi.org/10.1080/026999396380321

[22] M. E. Crosby and B. Auernheimer, "Physiological data feedback for application in distance education," in *In Proceedings of PUI 2001. NY: ACM. Pp*, 2001, pp. 1–5.

[23] "All about heart rate (pulse)." [Online]. Available: http://www.heart.org/HEARTORG/Conditions/HighBloodPressure/GettheFactsAboutHighBloodPressure/All-About-Heart-Rate-Pulse_UCM_438850_Article.jsp

[24] H. K. Walker, W. D. Hall, and J. W. Hurst, *Clinical methods: the history, physical and laboratory examinations.* Butterworths, 1990.

[25] "Before, during and after your emg test." [Online]. Available: https://www.hopkinsmedicine.org/healthlibrary/test_procedures/neurological/electromyography_92,P07656

[26] [Online]. Available: https://www.media.mit.edu/galvactivator/faq.html

[27] H. M. D. Morree, B. M. Szabó, G.-J. Rutten, and W. J. Kop, "Central nervous system involvement in the autonomic responses to psychological distress," *Netherlands Heart Journal*, vol. 21, no. 2, p. 64–69, 2012.

[28] D. Purves and S. M. Williams, *Neuroscience*. Sinauer Associates, 2001.

[29] "Ieee transactions on affective computing." [Online]. Available: https://ieeexplore.ieee.org/xpl/aboutJournal.jsp?punumber=5165369#AimsScope

[30] R. Picard, *Affective Computing*, ser. Inteligencia artificial. MIT Press, 2000. [Online]. Available: https://books.google.pt/books?id=GaVncRTcb1gC

[31] R. W. Picard *et al.*, "Affective computing," 1995.

[32] J. Tao and T. Tan, "Affective computing: A review," in *Affective Computing and Intelligent Interaction*, J. Tao, T. Tan, and R. W. Picard, Eds. Berlin, Heidelberg: Springer Berlin Heidelberg, 2005, pp. 981–995.

[33] P. N. Juslin and K. Scherer, *Vocal Expression of Affect*, 01 2005.

[34] "Toward Detecting Emotions in Spoken Dialogs," vol. 12, Mar. 2005, pp. 293–303. [Online]. Available: http://ict.usc.edu/pubs/Toward%20Detecting%20Emotions%20in%20Spoken%20Dialogs.pdf

[35] N. L. Etcoff and J. J. Magee, "Categorical perception of facial expressions," *Cognition*, vol. 44, no. 3, p. 227–240, 1992.

[36] P. Ekman and W. Friesen, *Facial Action Coding System*. Consulting Psychologists Press, 1978, no. v. 1. [Online]. Available: https://books.google.pt/books?id=pHZ3XwAACAAJ

[37] R. El Kaliouby and P. Robinson, "Real-time inference of complex mental states from facial expressions and head gestures," in *Real-time vision for human-computer interaction*. Springer, 2005, pp. 181–200.

[38] R. Calvo and S. K. D'Mello, "Affect detection: An interdisciplinary review of models, methods, and their applications," vol. 1, pp. 18–37, 01 2010.

[39] R. Walk and K. Walters, *Perception of the Smile and Other Emotions of the Body and Face at Different Distances*, 01 1988.

[40] H. Gunes and M. Piccardi, "Bi-modal emotion recognition from expressive face and body gestures," *Journal of Network and Computer Applications*, vol. 30, no. 4, pp. 1334 – 1345, 2007, special issue on Information technology. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S1084804506000774

[41] S. Poria, E. Cambria, A. Hussain, and G.-B. Huang, "Towards an intelligent framework for multimodal affective data analysis," *Neural Networks*, vol. 63, pp. 104 – 116, 2015. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0893608014002342

[42] W.-H. Lin, T. Wilson, J. Wiebe, and A. Hauptmann, "Which side are you on?: Identifying perspectives at the document and sentence levels," in *Proceedings of the Tenth Conference on Computational Natural Language Learning*, ser. CoNLL-X '06. Stroudsburg, PA, USA: Association for Computational Linguistics, 2006, pp. 109–116. [Online]. Available: http://dl.acm.org/citation.cfm?id=1596276.1596297

[43] C. Liu, K. Conn, N. Sarkar, and W. Stone, "Physiology-based affect recognition for computer-assisted intervention of children with autism spectrum disorder," *International Journal of Human-Computer Studies*, vol. 66, no. 9, pp. 662 – 677, 2008. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S1071581908000487

[44] N. Dael, M. Mortillaro, and K. R. Scherer, "Emotion expression in body action and posture." *Emotion*, vol. 12, no. 5, p. 1085, 2012.

[45] Y. Gu, X. Mai, and Y.-j. Luo, "Do bodily expressions compete with facial expressions? time course of integration of emotional signals from the face and the body," *PLoS One*, vol. 8, no. 7, p. e66762, 2013.

[46] D. Hung, T. Chau, and E. Biddiss, "Towards ambulatory emotion recognition," *CMBES Proceedings*, vol. 34, no. 1, 2018.

[47] E. L. Van Den Broek, V. Lisỳ, J. H. Janssen, J. H. Westerink, M. H. Schut, and K. Tuinenbreijer, "Affective man-machine interface: unveiling human emotions through biosignals," in *International Joint Conference on Biomedical Engineering Systems and Technologies*. Springer, 2009, pp. 21–47.

[48] C. Busso, Z. Deng, S. Yildirim, M. Bulut, C. M. Lee, A. Kazemzadeh, S. Lee, U. Neumann, and S. Narayanan, "Analysis of emotion recognition using facial expressions, speech and multimodal information," in *Proceedings of the 6th international conference on Multimodal interfaces*. ACM, 2004, pp. 205–211.

[49] B. Doerrfeld, "20 emotion recognition apis that will leave you impressed, and concerned | nordic apis |," Aug 2016. [Online]. Available: https://nordicapis.com/20-emotion-recognition-apis-that-will-leave-you-impressed-and-concerned/

[50] "Autism spectrum disorder - struggling with communication." [Online]. Available: https://www.psychologytoday.com/us/blog/you-say-more-you-think/201104/autism-spectrum-disorder-struggling-communication

[51] *Diagnostic and statistical manual of mental disorders: DSM-5.* American Psychiatric Association, 2013.

[52] "Autism spectrum disorder." [Online]. Available: https://www.nimh.nih.gov/health/topics/autism-spectrum-disorders-asd/index.shtml

[53] "Autism spectrum disorder (asd)," Apr 2018. [Online]. Available: https://www.cdc.gov/ncbddd/autism/data.html

[54] R. C. Network, "Emotional development in children with autism spectrum disorder." [Online]. Available: http://raisingchildren.net.au/articles/autism_spectrum_disorder_emotional_development.html

[55] C. Liu, K. Conn, N. Sarkar, and W. Stone, "Affect recognition in robot assisted rehabilitation of children with autism spectrum disorder," in *Proceedings 2007 IEEE International Conference on Robotics and Automation*, April 2007, pp. 1755–1760.

[56] H.-C. Chu, W. W.-J. Tsai, M.-J. Liao, and Y.-M. Chen, "Facial emotion recognition with transition detection for students with high-functioning autism in adaptive e-learning," *Soft Computing*, vol. 22, no. 9, pp. 2973–2999, May 2018. [Online]. Available: https://doi.org/10.1007/s00500-017-2549-z

[57] P. Washington, C. Voss, N. Haber, S. Tanaka, J. Daniels, C. Feinstein, T. Winograd, and D. Wall, "A wearable social interaction aid for children with autism," in *Proceedings of the 2016 CHI Conference Extended Abstracts on Human Factors in Computing Systems.* ACM, 2016, pp. 2348–2354.

[58] S. University. [Online]. Available: http://autismglass.stanford.edu/

[59] "Mp160 starter systems | biopac." [Online]. Available: https://www.biopac.com/product-category/research/systems/mp150-starter-systems/

[60] U. Dimberg, "Social fear and expressive reactions to social stimuli," *Scandinavian Journal of Psychology*, vol. 38, no. 3, p. 171–174, 1997.

[61] C. J. D. Luca, L. D. Gilmore, M. Kuznetsov, and S. H. Roy, "Filtering the surface emg signal: Movement artifact and baseline noise contamination," *Journal of Biomechanics*, vol. 43, no. 8, pp. 1573 – 1579, 2010. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0021929010000631

[62] A. A. Fedotov, "Selection of parameters of bandpass filtering of the ecg signal for heart rhythm monitoring systems," *Biomedical Engineering*, vol. 50, no. 2, pp. 114–118, Jul 2016. [Online]. Available: https://doi.org/10.1007/s10527-016-9600-8

[63] A. Greco, G. Valenza, and E. P. Scilingo, *Modeling for the Analysis of the EDA*. Cham: Springer International Publishing, 2016, pp. 19–33. [Online]. Available: https://doi.org/10.1007/978-3-319-46705-4_2

[64] L. Mason, *Signal processing methods for non-invasive respiration monitoring.* University of Oxford Oxford, 2002.

[65] J. C. Ives and J. K. Wigglesworth, "Sampling rate effects on surface emg timing and amplitude measures." *Clinical biomechanics*, vol. 18 6, pp. 543–52, 2003.

[66] "Rectified linear units (relu) in deep learning | kaggle." [Online]. Available: https://www.kaggle.com/dansbecker/rectified-linear-units-relu-in-deep-learning/code

[67] [Online]. Available: http://ufldl.stanford.edu/tutorial/supervised/SoftmaxRegression/

[68] M. Malik, "Heart rate variability: Standards of measurement, physiological interpretation, and clinical use," vol. 93, pp. 1043–1065, 03 1996.

[69] R. A. Barkley, "The ecological validity of laboratory and analogue assessment methods of adhd symptoms," *Journal of abnormal child psychology*, vol. 19, no. 2, pp. 149–178, 1991.

[70] [Online]. Available: http://www.vitaljacket.com/?lang=pt-pt

[71] "Biopac." [Online]. Available: https://www.biopac.com/corporate/about-biopac/

[72] "Understanding an ecg," Mar 2018. [Online]. Available: https://geekymedics.com/understanding-an-ecg/

[73] "Spearman's correlation." [Online]. Available: http://www.statstutor.ac.uk/resources/uploaded/spearmans.pdf

Appendices

Appendix A

Vital Jacket vs Biopac

The ecological validity of data collected in the laboratory is a concern [69], since the laboratory environment may induce alterations in natural response of the person. Therefore, it is interesting to find solutions able to transfer the data collection to real daily life routines. To accomplish such goal, it is necessary to validate wearable sensors against laboratory hardware. In this chapter, a discussion between the Vital Jacket [70] and the Biopac MP160 [71] is presented. Since, the Biopac [71] is the usual hardware on laboratory experiences, therefore it will be considered in this study as the ground-truth.

(a) Vital Jacket
(Source: `http://bit.ly/2IFQjFy`)

(b) Biopac MP160
(Source: `http://bit.ly/2Iaq1vo`)

Figure A.1: Different system used for this experiment

The objective of this chapter is to verify if using the biopac system as the ground truth they are equivalent.

To perform the comparison between Biopac and VJ signals, an experiment was designed at the Department of Education and Psychology of the University of Aveiro. [1]. During this experiment each of the participants used both the Vital Jacket and the Biopac sensors and watched three different types of emotion-inducing movies (Neutral, Fear and Happiness).

[1] `http://www.ua.pt/dep/`

A convenience sample was collected for this study, recruiting several students of multiple courses at the University of Aveiro. The participants were asked to fill a pre-selection questionnaire which addressed multiple variables, as: age; gender; health issues; actual medication (the questionnaire proposed can be found in: [2]). This questionnaire was also used to select candidates for a different experiment. The inclusion criteria comprised of the following: Portuguese nationality; age between 18 and 35 years old; having normal or corrected-to-normal vision. The exclusion criteria included: presence of a psychological/psychiatric disorder; having any other physical/health disorder which can impact the results (e.g., cardiac arrhythmia); not being undertaking any medication which impacts the data collection, or which can be indicative of a psychological/psychiatric disorder.

At this date the number of participants that were due to make the experience was 11 (8 females and 3 males), with an age mean of 20,73 and a standard deviation of 2.15 and their age varied between 18 and 34 years old, the males had a mean age of 21.67 with a standard deviation of 2.52 and the females had a mean age of 20.37 with a standard deviation of 2.06.

The data collection was divided into three different sessions per participant, each separated by a week, different emotion is elicited in each week.

A.1 Protocol and Setup

The experiment was performed in the Olfaction lab present at the Psychology and Education Department at the University of Aveiro [3]. The setup consisted of desktop computer, a chin rest, the sensors and a pair of headphones.

[2] `http://questionarios.ua.pt/index.php/331418?lang=pt`
[3] `http://www.ua.pt/dep/`

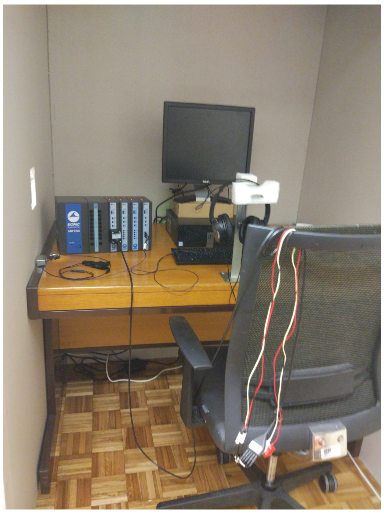


Figure A.2: Experimental setup at Olfaction Lab.

The desktop was the principal component, providing the videos that were going to be shown, as well as the questionnaires about their emotional state and videos. The mouse was positioned in a way that the user could easily interact with it without the need of fast movements. A chin rest was also present on the table, with a height of 32 cm and a distance to the screen that varies between 57 cm and 64 cm, it was positioned on the centre of the table, it was necessary to guarantee a constant distance between the participant and the centre of the screen.

Figure A.3: Closer setup perspective.

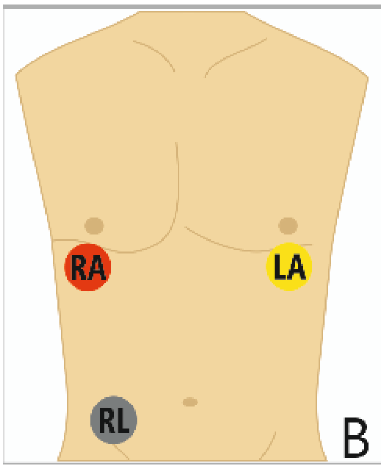
All of the participants were supplied with two sets of ECG electrodes, one of the sets of ECGs was from the Vital Jacket and the other ECG set. The positioning of them was accordingly to the VJ lead 1 guidelines as shown in the figure A.4a and the Biopac accordingly, to the lead 2 in the figure A.4b.

The electrodes positioning was as follow:

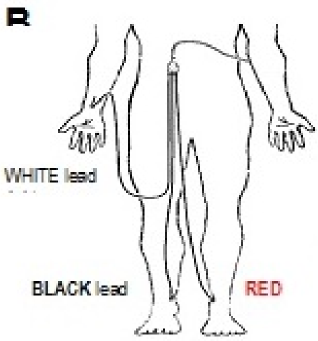
- Vital Jacket ECG - lead one positioning fig. A.4a. electrodes were positioned on the chest of the participants near the heart;
- Biopac ECG - lead two positioning fig. A.4b. Negative electrode was placed at the right arm near the palm. Positive electrode was placed at the left leg near the foot,

on the ankle. Ground electrode was placed at the right leg near the foot, on the ankle;

- Biopac EDA - positioned at the Medial Phalanx of the index and the middle finder of the left hand A.4c;



(a) Vital Jacket 1l electrode positioning
(Source: `https://goo.gl/Pt2xAS`)



(b) Biopac sensors positions
(Source: `https://goo.gl/VohbtS`)

(c) EDA electrodes positioning
(Source: `https://goo.gl/VohbtS`)

Figure A.4: Electrodes positioning

Each of the participants were subject to three different phases of the experiment, each one of them with a week of separation. The order of the sessions was counterbalanced, that was, the same number of participants executed the experiment in a same specific order. This was done so the variable of the order couldn't be a co-variable of the data.

The selection of the videos was done comparing two pilot tests done in previous experiments, the two were compared and the one with the results that were more interesting to this study was selected.

The protocol executed for each of the phases was the same, the only variable that change was the set of the videos watched during the video segment part, depending on the emotion to be induced.

An android application was also created to perform the synchronisation of the signals, this application connects to the Vital Jacket sensor through Bluetooth and two triggers can be manually set to indicate the beginning and end of each section. The data at the end would be sent to the cloud.

The protocol of the experience, containing all three phases is described next:

- With the selection of the participants done, a first interview was scheduled, to explain the procedure. At the same time, the written consent of the participant was also collected. They were also informed that they could quit from the experiment at any time, if they feel the necessity to do so. All the data was told to be private and only been used for the presented objective. Then, three sessions were scheduled according with the availability of the experimenter/participant, the participants were than told they would watch a series of videos (35-40 minutes) per session while their psychophysiological response would be also recorded. The data collection took place between May and June of 2018.

- At the beginning of each of the 3 experiments these steps where always done: they were asked to sit on a chair and answer three different questionnaires - STICSA State, SWLS [4] and PANAS [5]. In the second and third session they were asked to fill the STICSA-State only. After this step, the experimenter asked the participant to get up and cleaned with ethyl alcohol the areas in which the vital jacket electrodes should be placed and then the electrodes were carefully placed. Then the experimenter asked the participant to sit up in a chair next to the computer and the same procedure was followed for the placement of BIOPAC electrodes.

- After all the electrodes were placed it was made a resting phase so the signals would stabilise (approximately 10 minutes).

- Then, the experimenter recalled all the procedure with the participant, stressing that it was important to try to stay as quiet as possible and that if they have any question they should clarify it. After making sure that the participant understood all the procedure, his/her right to withdraw from the experiment at any time and that he/she was comfortable and ready to start, the experiment, several questions about the level of anxiety, fear, joy and stress, appeared on the monitor, and the participant should answer via an analogy visual scale which varied between ”Nothing” and ”Much”. The participants were told they should select the point of the scale that best represented their answer with the mouse and the values are converted from 0 to 100. After all the questions were answered, a page with the next procedure was

[4] `https://www.psytoolkit.org/survey-library/satisfaction-with-life.html`

[5] `https://www.psytoolkit.org/survey-library/affect-panas.html`

shown saying that they would watch next the first video and that they should not move, so the signal wouldn't get noise from the movement.

- All the triggers were deployed manually, so, at the beginning and end of each segmentation (video baseline vs. group of emotional videos) the experimenter would click almost synchronously on both the mobile device that contained the app that was connected to the Vital Jacket and on a button on the Keyboard (F5, F6, F7, F8) for the Biopac software running. These triggers were then used to synchronize the signals.

- The first video shown (baseline video) was a segment from a documentary, and there were 3 different baseline videos with 5 minutes approximately each.

- After the conclusion of the baseline video, some questions appeared in the screen. These questions addressed how the user rated the video and what he/she felt during the video, in terms of arousal, intensity, pleasantness and emotions as happiness, fear and anxiety. These questions were about how the user felt and how the user thought the other participants would have felt if they watched the video, in order to compare the ratings. The order of the questions (how they felt vs. how the others would have felt) was counterbalanced.

- After this baseline video evaluation, the experiment proceeded and the emotional videos were presented. The emotional videos were a set of videos for each emotion (12 for fear with a total duration of 33.55 minutes, 8 for happiness with a total duration of 29.32 minutes and 12 neutral videos with a total duration of 34.31 minutes), that were presented with an interval of 1 second between segments. The emotional videos were always different in the three sessions (in one session were presented segments of comedy videos, in another session were presented segments of terror movies and in the third session were presented segments of documentaries - the order of the emotional condition was counterbalanced).

- After the emotional videos ended, the same questions about how the user and the others would have felt during the videos, appeared. Then, the questions made on the start of the experiment was also repeated (about current emotional state).

- At the end of the third session each user would also fill two additional questionnaires, namely the MAIA (body consciousness) [6] and LEAS-A (levels of emotional awareness). Lastly, the participants were debriefed and any question about the experiment was clarified. The participants were told that in the end of the data collection a lottery would be made to thank them for their time and collaboration.

[6] `https://goo.gl/E3yNnL`

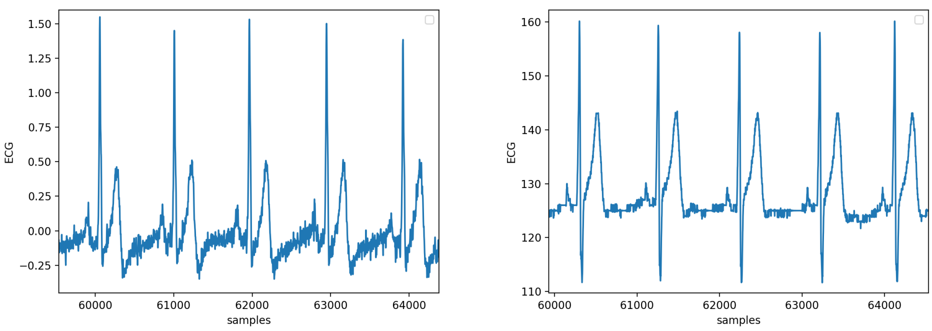
Figure A.5: Experiment setup example

With the experiment done a data set was created, this data set was constituted of 6 different files per participant, this where 2 different files per session (one file which will represent the ECG from the Vital Jacket sensor and the other file that will have the recorded ECG from the Biopac system) for all the 3 conditions.

A.2 Methods

So, the signal would correspond to the same time window, and correspond to the time window of the experiment a cut was done based on the triggered deployed by both the android application to the Vital Jacket and the software on the desktop to the Biopac. With this cut done we knew that signal would correspond to the experiment in each of the variables.

An example of both signals in the same time window can be analysed in the figure A.6a for the Biopac signal and A.6b for the Vital Jacket signal.



(a) Recorded ECG from Biopac (b) Recorded ECG from Vital Jacket

Figure A.6: Recorded ECG from both sensors

Two different ECG leads were used, the Vital Jacket use the first limb lead, as for the Biopac it was used the second limb lead (the reason for this choice was because there was also an EDA sensor on the left hand of the participant, and using the first limb lead and putting an ECG sensor on the same hand could cause interference), the problem with this was that the ECG produced by the Biopac had a R-peak with a higher amplitude, this was caused because this is the lead that is more correlated with the direction of the electrical spread [72].

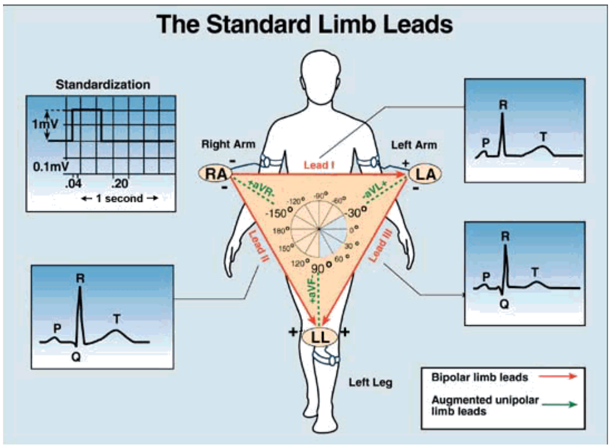


Figure A.7: The standard limb leads
(source: `https://ecg.utah.edu/lesson/1`)

The file recorded by the Vital Jacket was a text file that had the timestamp, a vector of ECG associated with that timestamp and the ID of the stimulus (0 - Nothing happening, 1 - Baseline running or 2 - Movie section in progress). As for the Biopac file it was recorded in an ACQ file, this file was easily converted to MAT file with the Bioread library [7], the stimulus of this file were separated into a different file that had simply the time and the name of the stimulus.

As the sampling rate used for this experiment was of 500Hz for the Vital Jacket and of 1000 Hz for the Biopac, some kind of correction was necessary to both of the signals could be correctly compared. Considering that the Biopac signal is our ground truth, the ECG from the Vital Jacket was up sampled to 1000 Hz.

To proceed with this up sampling it was used the resample method [8] from the *scipy library* , this method takes a signal with x samples and turns into a signal with xn samples, being the xn a number defined by the user, this transformation was based on the Fourier method.

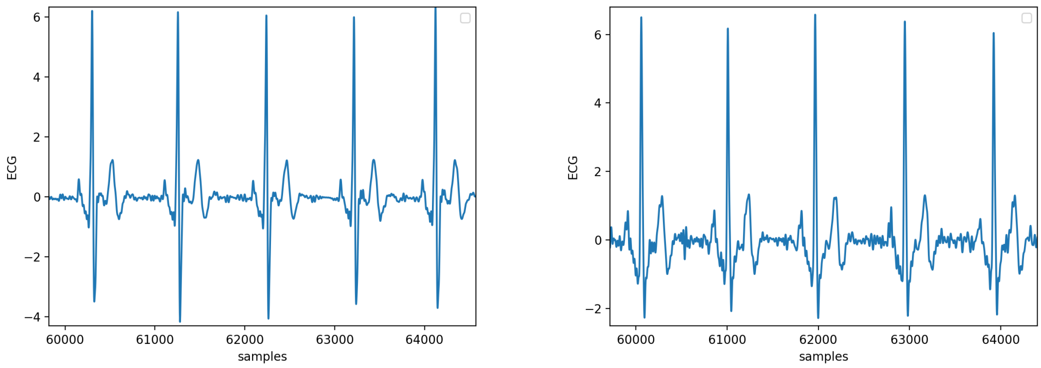
After the up sampling each of the signals was passed through a filter provided by the Neurokit library [9] and the returned signal was analysed.

[7]`https://github.com/njvack/bioread`

[8]`https://docs.scipy.org/doc/scipy/reference/generated/scipy.signal.resample.html`

[9]`http://neurokit.readthedocs.io/en/latest/documentation.html#emg-process`

The filter used on this signal was the same as described in section 4.3 and the resulting filtered signal can be verified in fig. A.8a and A.8b.



(a) Filtered ECG from Vital Jacket

(b) Filtered ECG from Biopac

Figure A.8: Filtered ECG from both sensors

In order to compare the two ECG signals, the data was normalised. The two hardware systems have a different resolution, so they vary in different values.

The R-peaks were identified, and their indexes were used to try to overlap the two signals. Since, different lead configuration, hardware and applications were used, there were delays that need to be corrected to be possible compare the two signals. There were two possible solutions to this, the first was to calculate the delta (difference) between each of the r-peaks:

$$\Delta = ApplicationDelay + LeadDelay + \varepsilon$$

Being the ε a random value that is produced by the sensor. It was used the mean value of the difference between each of the r-peaks:

$$\bar{\Delta} = (\Delta_{1VJ} - \Delta_{1B}) + ... + (\Delta_{NVJ} - \Delta_{NB})$$

After all the pre-processing was done the comparison phase could start.

The comparison was evaluated in both the complete signals and in segments of 30 seconds and 5 seconds. Three different measures were calculated to compare both the signals:

- Spearman's rank-order correlation - it measures the strength and direction of association between the two different signals;
- Mutual Information - this is the measure of the mutual dependence between the two variables, in other words, it assesses the "amount of information" obtained about one of the signals, throughout the other one. The maximizing for the mutual information was calculated by comparing two equal signs from Biopac, and it had the value of 7.73 for the ECG comparison and 7.75 for the Heart Rate comparison;

- Euclidean Distance - is the distance between straight line between two points in the Euclidean space;

A.3 Comparison

The R-Peaks were the feature decided to use for the synchronisation because is one of the features that is most visible and more easily differentiated from the other, on contrary of the Q, P, S, T Waves.

As one of the metrics used for the comparison was the Mutual information the maximising value needed to be calculated, for this it was done the mutual information between the same Biopac Signal and the value 7.73 was acquired, so, the closer to this value the better the results are.

For the Euclidean distance the minimising was 0, so, the closer to this value the better the results are.

For the Spearman correlation rank the table A.1 was taken into consideration [73].

Table A.1: Spearman rank correlation value range and their meaning

Spearman rank correlation value range	Correlation meaning
0.00 - 0.19	"very weak"
0.20 - 0.39	"weak"
0.40 - 0.59	"moderate"
0.60 - 0.79	"strong"
0.80 - 1.00	"very strong"

(Source: https://goo.gl/54DPQi)

Before synchronisation, the signals representation is presented in Figure: A.9.

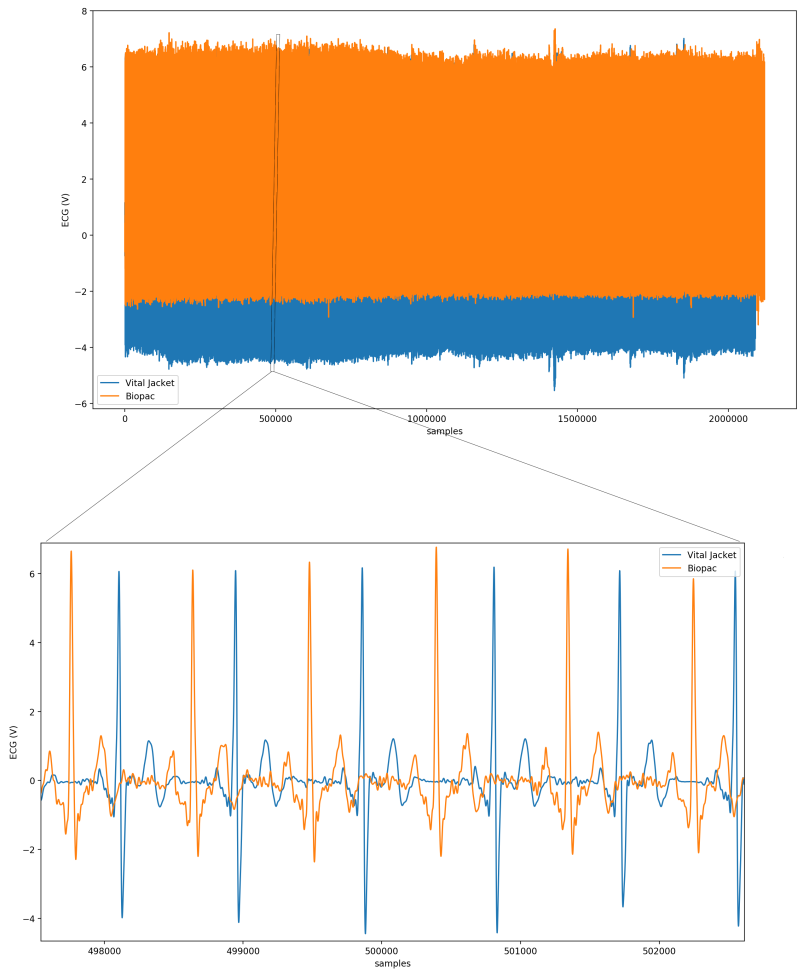


Figure A.9: Original complete signal comparison

A synchronisation was then executed using the mean difference between the R-Peak from the same beat values (Figure: A.10).

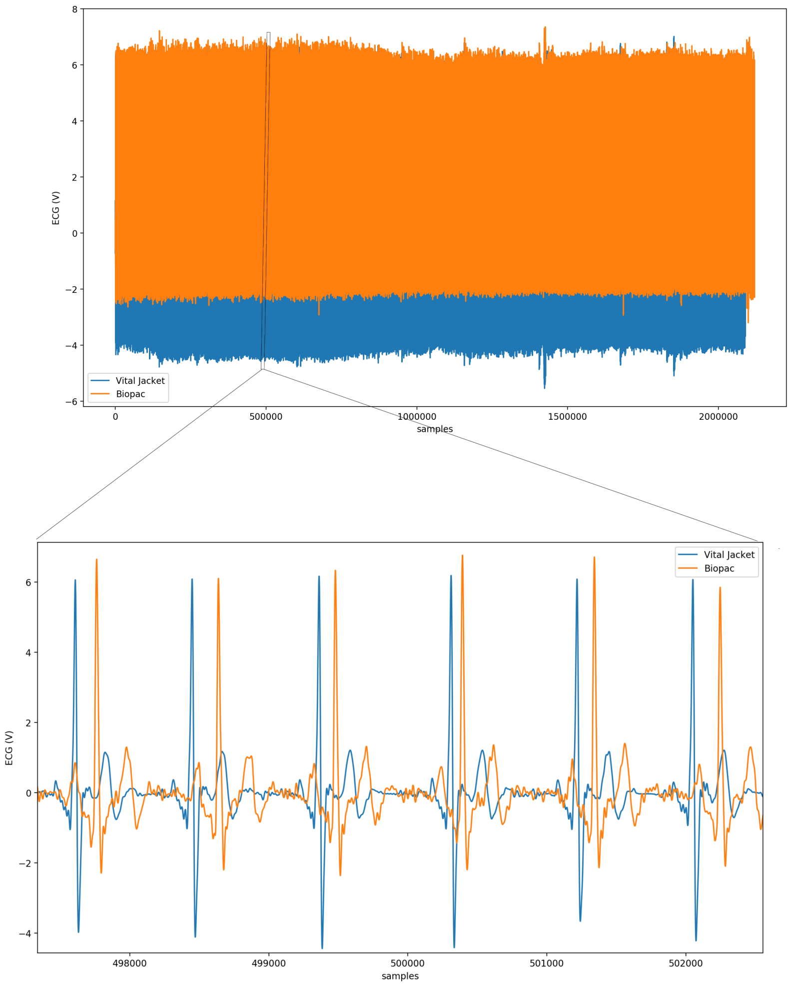


Figure A.10: Original complete signal comparison with mean correction

Even with this done there wasn't a full synchronisation of the signals, this could be caused by several unknown delays.

The comparison measures are presented in Table: A.2, the standard deviation is always zero as this is a unique value.

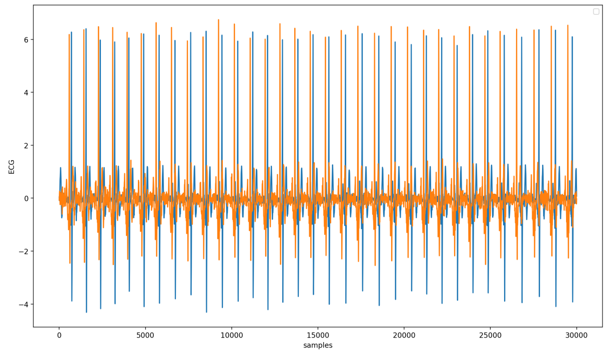
Table A.2: Metrics from the full time experiment

Metric	Value	Standard Deviation
Spearman's rank correlation	-0.01	0.00
Euclidean distance	201	0
Mutual Information	142	0

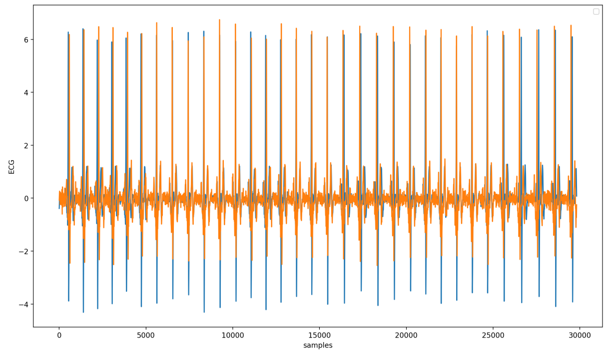
With a "very weak" [73] Spearman correlation, a Euclidean distance with a distance of 201 from his minimum value and a mutual information with a difference of 135 values from his maximum, it could be concluded that the ECG from the Vital Jacket was not comparable with the Biopac ECG. The main reason for these values was the random delay inserted in the ECG that does not allow a true synchronisation neither comparison of the signals.

So, to try and fight this results, each of the signals was divided into sectors of 30 seconds (Figure: A.11a) and 5 seconds, it was synchronised in two different ways to check which one was the better to use, the first one was with the mean difference between the R-Peaks (Figure: A.11b) and a second one (just for the 30 seconds case, as the other has in the other it was not possible) that only used the mean difference between the first six R-Peaks. Six R-peaks were selected since by visual inspection of the 30 seconds ECG segments a pattern was found, the first six values were almost at the same distance between them (Figure: A.11c).

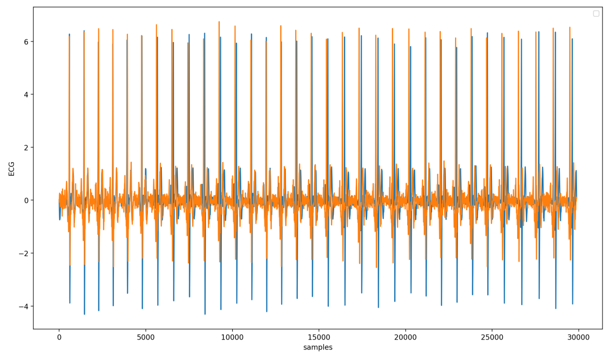
As there were done rigid cuts to the ECG to create the windows some parts of it due to the delay could pass to the next window while the corresponding peak from the other signal remained on the previous window. To try to solve this, an overlap window was used. A 50% overlap window was used in the two cases (15 seconds for the 30 seconds part and 2.5 seconds for the 5 seconds) (Figure: A.4, A.4 and Table: A.13, A.13).



(a) Comparison of both signals, not synchronised with a cut of 30 seconds



(b) Comparison of both signals, synchronised with the mean difference between the R-Peaks in a cut of 30 seconds

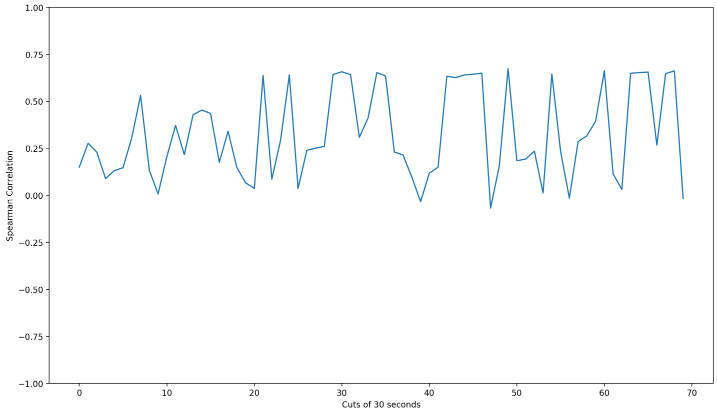


(c) Comparison of both signals, synchronised with the mean difference between the first six R-Peaks in a cut of 30 seconds

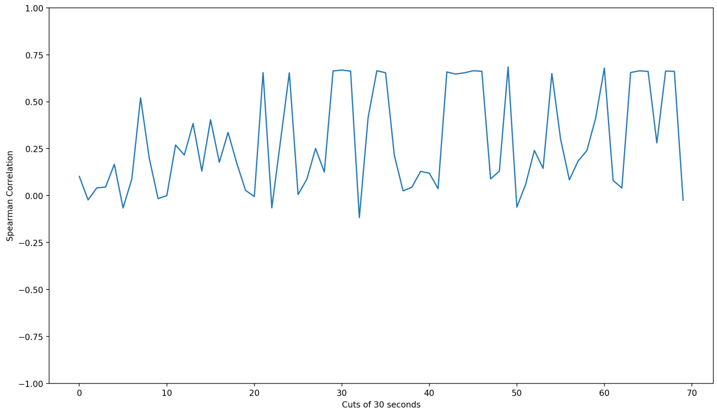
Figure A.11: Comparison of both signals, with different synchronisations in a cut of 30 seconds

The representation on (Figure: A.11c is replicated for the other ECG segments. So, it can be extrapolated that both of the signals look similarly synchronised, the difference starts to increase towards the ending (Figure: A.11b). For a deeper analysis of this synchronisation the metrics described in (Section: 5.3) are going to be used and compared. The Spearman correlation time series is presented for both of the cases (Figure: A.12a for mean difference of all R-Peaks and Figure: A.12b for the mean difference of the first six R-Peaks), the mean of this coefficient as the other metrics were also calculated (Table: A.3 for mean difference of all R-Peaks and Figure: A.4 for the mean difference of the first six R-Peaks).

As it can be checked on the tables A.3 and A.4, the Euclidean Distance and Mutual Information values are almost the same, so no deduction from here can be taken. The real difference lies on the Spearman rank correlation, from here it can be seen that the value from the mean difference between all the R-Peaks is greater, even if not by a large margin. Some negative Spearman rank correlation values can be seen on the figures A.12a and A.12b, this is caused by the random delays inserted what causes the signals to switch places.



(a) Spearman rank correlation time series from the comparison of both signals, synchronised with the mean difference between the R-Peaks in a cut of 30 seconds



(b) Spearman rank correlation time series from the comparison of both signals, synchronised with the mean difference between the first six R-Peaks in a cut of 30 seconds

Table A.3: Metrics from the comparison of both signals, synchronised with the mean difference between all the R-Peaks in a cut of 30 seconds

Metric	Value	Standard Deviation
Spearman's rank correlation	0.31	0.23
Euclidean distance	198	56
Mutual Information	10	0.05

Table A.4: Metrics from the comparison of both signals, synchronised with the mean difference between the first six R-Peaks in a cut of 30 seconds

Metric	Value	Standard Deviation
Spearman's rank correlation	0.28	0.27
Euclidean distance	199	65
Mutual Information	10	0.05

The results with the overlap window of 15 seconds can be verified in the table A.5. Compared with the results from tables A.3 and A.4, the values from the Euclidean distance and the Mutual Information have a little change, again, the greater improvement was on the Spearman rank correlation, even if not for a great margin again.

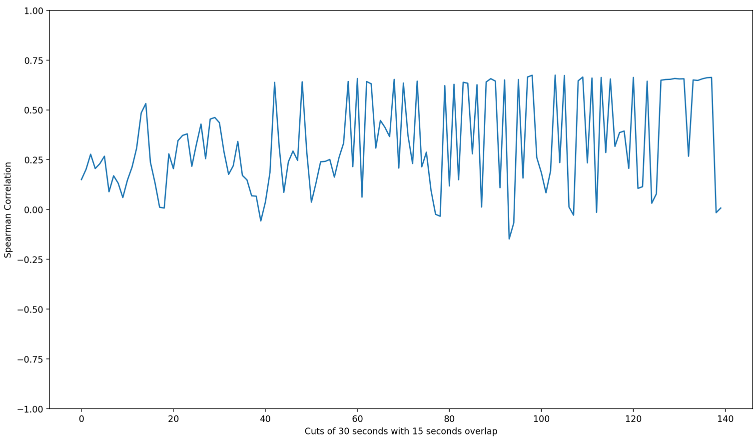


Figure A.13: Spearman rank correlation time series from the comparison of both signals, synchronised with the mean difference between the R-Peaks in a cut of 30 seconds with a overlap window of 15 seconds

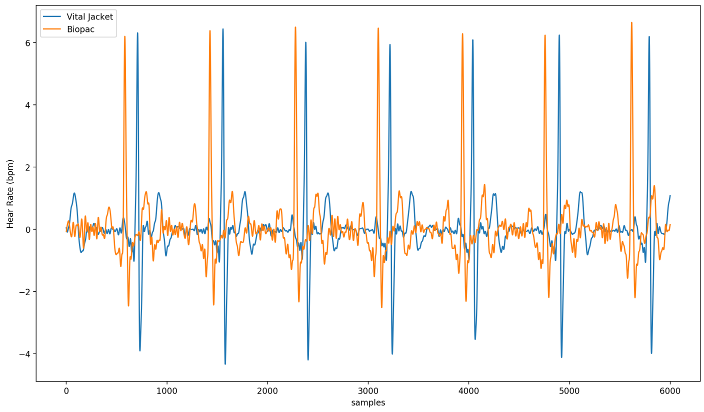
Table A.5: Metrics from the comparison of both signals, synchronised with the mean difference between the first six R-Peaks in a cut of 30 seconds with overlap window of 15 seconds

Metric	Value	Standard Deviation
Spearman's rank correlation	0.33	0.23
Euclidean distance	191	58
Mutual Information	10	0.17

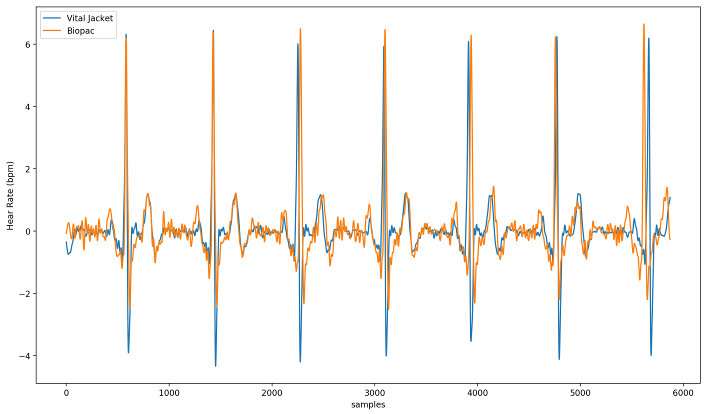
After analysing the results from the tables A.3, A.4 and A.5 it can be concluded that the best results are going to be gotten when using an overlap window with the mean difference between all R-Peaks for the synchronisation. But with values below 0.5 for the Spearman correlation it can be seen as a ”weak” correlation [73], for the values of Euclidean Distance all of them have a difference greater than 190 of his minimise, and for the Mutual information with a difference of 2.27 from his maximise, it can be concluded that they almost represent the same information.

So, with this it can be concluded that the values are not comparable, but this results can be caused by the synchronisation.

By results inspection, several areas from the signal were not synchronised, but it can also be observed that there is a pattern in the number of peaks synchronised. So, the window may be reduced, and the synchronisation can be improved obtaining a better performance in the metrics. So, a new window was created with a size of 5 seconds, it was only done the synchronisation from the mean difference of R-Peaks as the other synchronisation would be impracticable here, and the results are analysed again.



(a) Comparison of both signals, not synchronised with a cut of 5 seconds



(b) Comparison of both signals, synchronised with the mean difference between the R-Peaks in a cut of 5 seconds

Figure A.14: Comparison of both signals, with different synchronisations in a cut of 5 seconds

As it can be checked on the table A.6, compared with the results from the previous cut (table A.3, A.4 and A.5, with a 5 second cut and 30 second cut), both the Euclidean distance value and the Spearman rank correlation greatly improved, because of the value represented by the Spearman correlation between the signals can now be said as ”moderate” [73]. For the value of Euclidean Distance the difference is 78 of his minimise, being that they are much closer from being in top of each other (being similar) than before. For the Mutual information with a difference not greater than 1 from his maximise for all of them, it can be concluded that they almost represent the same information.

With this it can be concluded that with a lower window, a greater synchronisation is possible to achieve.

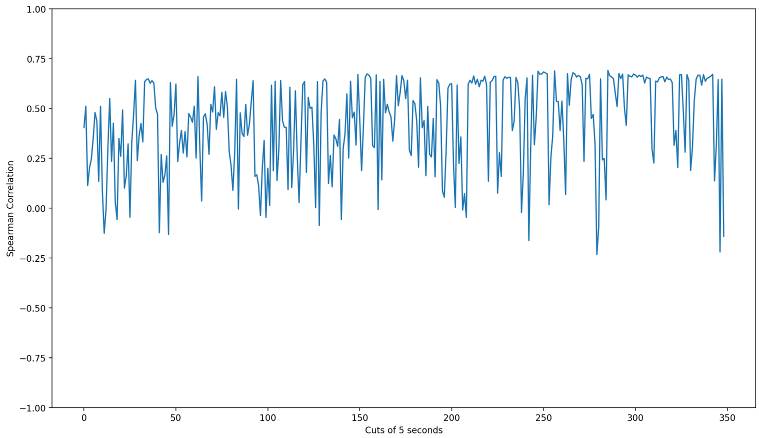


Figure A.15: Spearman rank correlation time series from the comparison of both signals, synchronised with the mean difference between the R-Peaks in a cut of 5 seconds

Table A.6: Metrics from the comparison of both signals, synchronised with the mean difference between the first six R-Peaks in a cut of 5 seconds

Metric	Value	Standard Deviation
Spearman's rank correlation	0.43	0.22
Euclidean distance	78	26
Mutual Information	8.61	0.08

The results with the overlap window of 2.5 seconds can be verified in the table A.7, compared with the results from tables A.6 all the values have an increase in their values. However, that improvement is too low. Concluding that an overlap window of this value does not give significant advantages, and maybe the computational power to do so is not worthy it.

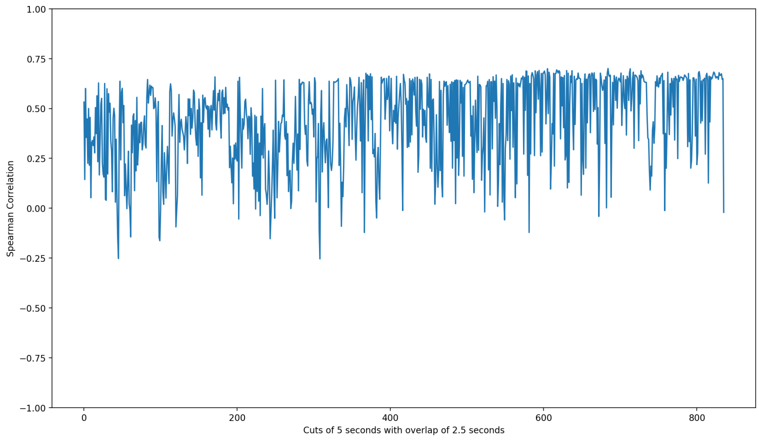


Figure A.16: Spearman rank correlation time series from the comparison of both signals, synchronised with the mean difference between the R-Peaks in a cut of 5 seconds with a overlap window of 2.5 seconds

Table A.7: Metrics from the comparison of both signals, synchronised with the mean difference between the first six R-Peaks in a cut of 5 seconds with a overlap window of 2.5 seconds

Metric	Value	Standard Deviation
Spearman's rank correlation	0.44	0.21
Euclidean distance	72	23
Mutual Information	8.42	0.54

After analysing the results, it can be concluded that the best results are going to be acquired with a lower window, this because the synchronisation gotten is better, this can be verified on the values present on the tables A.3, A.4, A.5, A.6 and A.7. With a window as low as 5 seconds a "moderate" correlation [73] can be achieved, a Euclidean Distance presents a low value, and the Mutual information is closer to the maximum.It can be concluded that the two signals almost represent the same information, concluding that they are more comparable than before.

Heart Rate Comparison

Usually, several measures were extracted from the ECG to analyse emotional and physiological context. So, thinking in the scenario to move from the lab to the real-life scenario, the usual metrics should also be validated. The heart rate (HR) is an example, and it is analysed, by the implementation of the same metrics.

The maximise of the Mutual Information is now of 7.75 calculated with the comparison between the same Hear Rate from the Biopac.

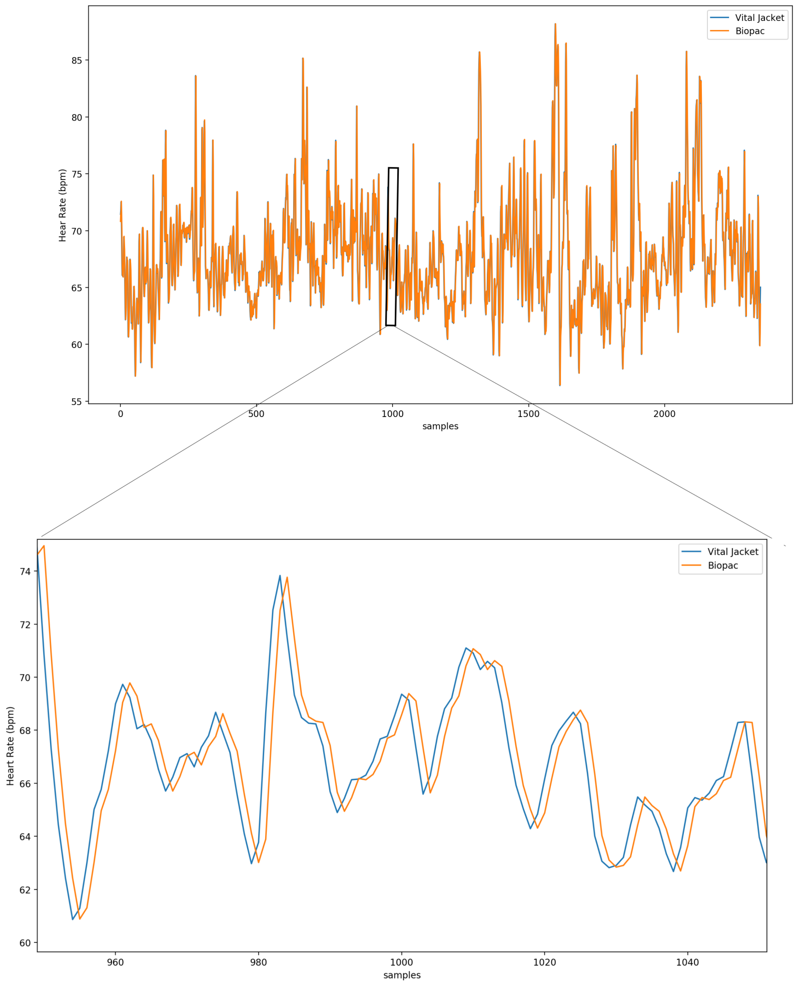


Figure A.17: Comparison of both Heart Rates, not synchronised

Table A.8: Metrics from the comparison of both Heart Rates, not synchronised

Metric	Mean	Standard Deviation
Spearman's rank correlation	0.92	0.00
Euclidean distance	80	0
Mutual Information	7.7	0.0

The results from the not synchronised comparison are already pretty good, as it can be seen on the table A.9 the Spearman correlation achieves a "very strong" correlation [73]. The Euclidean value is high because of the delay present on one of the signals as it can be seen on figure A.18, being comparable with the Euclidean distance between the ECG signals when this are synchronised with a low window. The mutual information is almost the same with a difference of 0.05, meaning that the signals represent almost the same information.

This indicated that the HR extracted from both system is comparable, and therefore the VJ system can be used in real life conditions to interpret the emotional response. However, in the next step the delay between both HR was corrected and the results again evaluated, this is shown on table A.9. This synchronisation was possible using the difference between the global maximum of both signals. There was an evolution on the Spearman value, with a score almost equal to one, being that was a ”very strong” correlation [73], the Euclidean distance present here is the lower gotten in all comparison with a value much closer to his minimise (0), that is, they are almost in top of each other, the mutual information has a difference of 0.33, meaning that they almost have the same information present.

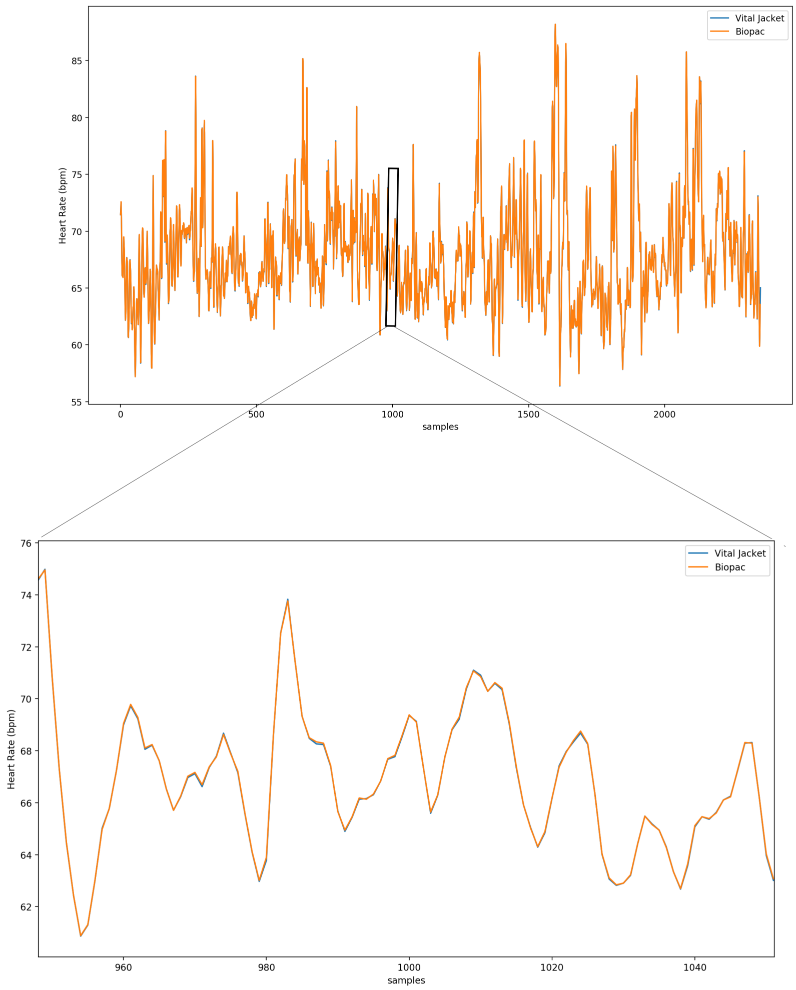


Figure A.18: Comparison of both Heart Rates, synchronised with difference between the global maximum of the Biopac Signal

Table A.9: Metrics from the comparison of both Heart Rates, synchronised with difference between global maximum

Metric	Mean	Standard Deviation
Spearman's rank correlation	0.99	0.00
Euclidean distance	1.91	0.00
Mutual Information	7.42	0.00

A.4 Results summary

Concluding, time series extracted from the ECG reveal to be almost similar between the two systems. Nevertheless, if the ECG is to be used, careful should be taken in order to deal with the delay observed in the VJ system. So, it can be concluded that the VJ system is a good choice to be used in daily routine data collection.

The table A.10 shows the mean values and standard deviation of all the metrics when the signal from all the experiences are putted to the test.

Table A.10: Metrics from the mean of multiple comparisons of both Heart Rates and ECG, synchronized

Metric	Spearman rank Mean	Spearman rank std	Euclidean Mean	Euclidean std	Mutual Information Mean	Mutual Information std
Heart Rate Comparison	0.96	0.06	46.11	7.49	7.73	0.09
ECG Comparison full signal	0.11	0.12	1983.94	60.18	14.51	0.06
ECG Comparison 30 second window	0.29	0.21	221.57	42.68	10.29	0.06
ECG Comparison 30 second window with overlap 15 second	0.30	0.21	217.73	47.78	10.28	0.15
ECG Comparison 5 second window	0.43	0.23	22.79	80.71	8.46	0.07
ECG Comparison 5 second window with overlap 2.5 second	0.43	0.23	80.43	22.99	8.46	0.06